





## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 15/31, 1/21, C12P 21/02, C07K 14/33, A61K 38/16, 39/08

(11) International Publication Number:

WO 98/07864

(43) International Publication Date:

26 February 1998 (26.02.98)

(21) International Application Number:

PCT/GB97/02273

A1

(22) International Filing Date:

22 August 1997 (22.08.97)

(30) Priority Data:

9617671.4 9625996.5

23 August 1996 (23.08.96)

GB 13 December 1996 (13.12.96) GB

(74) Agent: SCHLICH, George, William; Mathys & Squire, 100 Gray's Inn Road, London WCIX 8AL (GB).

Conrad, Padraig [GB/GB]; Microbiological Research

Authority CAMR (Centre for Applied Microbiology & Research), Porton Down, Salisbury, Wiltshire SP4 OJG (GB).

FOSTER, Keith, Alan [GB/GB]; Microbiological Research Authority CAMR (Centre for Applied Microbiology & Research), Porton Down, Salisbury, Wiltshire SP4 OJG

(60) Parent Application or Grant

(63) Related by Continuation

US 08/782,893 (CIP) Filed on 27 December 1996 (27.12.96)

(71) Applicants (for all designated States except US): MICRO-BIOLOGICAL RESEARCH AUTHORITY CAMR (CEN-TRE FOR APPLIED MICROBIOLOGY & RESEARCH) [GB/GB]; Porton Down, Salisbury, Wiltshire SP4 0JG (GB). THE SPEYWOOD LABORATORY LIMITED [GB/GB]; 14 Kensington Square, London W8 5HH (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): SHONE, Clifford, Charles [GB/GB]; Microbiological Research Authority CAMR (Centre for Applied Microbiology & Research), Porton Down, Salisbury, Wiltshire SP4 0JG (GB). QUINN, (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD. RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD,

**Published** 

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: RECOMBINANT TOXIN FRAGMENTS

(57) Abstract

A polypeptide has first and second domains which enable the polypeptide to be translocated into a target cell or which increase the solubility of the polypeptide, or both, and further enable the polypeptide to cleave one or more vesicle or plasma-membrane associated proteins essential to exocytosis. The polypeptide thus combines useful properties of a clostridial toxin, such as a botulinum or tetanus toxin, without the toxicity associated with the natural molecule. The polypeptide can also contain a third domain that targets it to a specific cell, rendering the polypeptide useful in inhibition of exocytosis in target cells. Fusion proteins comprising the polypeptide, nucleic acids encoding the polypeptide and methods of making the polypeptide are also provided. Controlled activation of the polypeptide is possible and the polypeptide can be incorporated into vaccines and toxin assays,

# FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AL AM AT AU AZ BA BB BF BG BJ BR CA CF CG CH CU CD DE DK EE | Afbania Armenia Austria Austria Australia Azerbaijan Bosnia and Herzegovina Barbados Belgium Burkina Faso Bulgaria Benin Brazil Belarus Canada Central African Republic Congo Switzerland Côte d'Ivoire Cameroon China Cuba Czech Republic Germany Denmark Estonia | ES FI FR GA GB GE GH GN GR HU IE IL IS IT JP KE KC KP KR LC LL LR | Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel Iceland Italy Japan Kenya Kyrgyzstan Democratic People's Republic of Korea Republic of Korea Republic of Korea Kazakstan Saint Lucia Liechtenstein Sri Lanka Liberia | LS I.T LU I.V MC MD MG MK ML MN MX NE NL NO NZ PL PT RO RU SD SE SG | Lesotho Lithuania Laxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia Mauritania Malawi Mexico Niger Netherlands Norway New Zealand Poland Portugal Romania Russian Pederation Sudan Sweden Singapore | SI<br>SK<br>SN<br>SZ<br>TD<br>TG<br>TJ<br>TM<br>TR<br>TT<br>UA<br>UG<br>US<br>UZ<br>VN<br>YU<br>ZW | Slovakia Senegal Swaziland Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago Ukraine Uganda United States of America Uzbekistan Viet Nam Yugoslavia Zimbabwe |
|---|--|---|---|---|---|--|---|
|---|--|---|---|---|---|--|---|

WO 98/07864 PCT/GB97/02273

### RECOMBINANT TOXIN FRAGMENTS

This invention relates to recombinant toxin fragments, to DNA encoding these fragments and to their uses such as in a vaccine and for *in vitro* and *in vivo* purposes.

The clostridial neurotoxins are potent inhibitors of calcium-dependent neurotransmitter secretion in neuronal cells. They are currently considered to mediate this activity through a specific endoproteolytic cleavage of at least one of three vesicle or pre-synaptic membrane associated proteins VAMP, syntaxin or SNAP-25 which are central to the vesicle docking and membrane fusion events of neurotransmitter secretion. The neuronal cell targeting of tetanus and botulinum neurotoxins is considered to be a receptor mediated event following which the toxins become internalised and subsequently traffic to the appropriate intracellular compartment where they effect their endopeptidase activity.

The clostridial neurotoxins share a common architecture of a catalytic L-chain (LC, ca 50 kDa) disulphide linked to a receptor binding and translocating H-chain (HC, ca 100 kDa). The HC polypeptide is considered to comprise all or part of two distinct functional domains. The carboxy-terminal half of the HC (ca 50 kDa), termed the  $H_{\rm c}$  domain, is involved in the high affinity, neurospecific binding of the neurotoxin to cell surface receptors on the target neuron, whilst the amino-terminal half, termed the  $H_{\rm N}$  domain (ca 50 kDa), is considered to mediate the translocation of at least some portion of the neurotoxin across cellular membranes such that the functional activity of the LC is expressed within the target cell. The  $H_{\rm N}$  domain also has the property, under conditions of low pH, of forming ion-permeable channels in lipid membranes, this may in some manner relate to its translocation function.

For botulinum neurotoxin type A (BoNT/A) these domains are considered to reside within amino acid residues 872-1296 for the  $H_{\rm C}$ , amino acid residues 449-871 for the  $H_{\rm N}$  and residues 1-448 for the LC. Digestion with trypsin effectively degrades the  $H_{\rm C}$  domain of the BoNT/A to generate a non-toxic fragment designated  $LH_{\rm N}$ ,

which is no longer able to bind to and enter neurons (Fig. 1). The  $LH_N$  fragm int so produced also has the property of enhanced solubility compared to both the par int holotoxin and the isolated LC.

It is therefore possible to provide functional definitions of the domains within the neurotoxin molecule, as follows:

- (A) clostridial neurotoxin light chain:
- -a metalloprotease exhibiting high substrate specificity for vesicle and/or plasma membrane associated proteins involved in the exocytotic process. In particular, it cleaves one or more of SNAP-25, VAMP (synaptobrevin / cellubrevin) and syntaxin.
- (B) clostridial neurotoxin heavy chain H<sub>N</sub> domain:
- -a portion of the heavy chain which enables translocation of that portion of the neurotoxin molecule such that a functional expression of light chain activity occurs within a target cell.
- -the domain responsible for translocation of the endopeptidase activity, following binding of neurotoxin to its specific cell surface receptor via the binding domain, into the target cell.
- -the domain responsible for formation of ion-permeable pores in lipid membranes under conditions of low pH.
- -the domain responsible for increasing the solubility of the entire polypeptide compared to the solubility of light chain alone.
- (C) clostridial neurotoxin heavy chain H<sub>c</sub> domain.
- -a portion of the h avy chain which is responsible for binding of the native

holotoxin to cell surface receptor(s) involved in the intoxicating action of cl stridial t xin prior to internalisation of the toxin into the cell.

The identity of the cellular recognition markers for these toxins is currently not understood and no specific receptor species have yet been identified although Kozaki et al. have reported that synaptotagmin may be the receptor for botulinum neurotoxin type B. It is probable that each of the neurotoxins has a different receptor.

It is desirable to have positive controls for toxin assays, to develop clostridial toxin vaccines and to develop therapeutic agents incorporating desirable properties of clostridial toxin.

However, due to its extreme toxicity, the handling of native toxin is hazardous.

The present invention seeks to overcome or at least ameliorate problems associated with production and handling of clostridial toxin.

Accordingly, the invention provides a polypeptide comprising first and second domains, wherein said first domain is adapted to cleave one or more vesicle or plasma-membrane associated proteins essential to neuronal exocytosis and wherein said second domain is adapted (i) to translocate the polypeptide into the cell or (ii) to increase the solubility of the polypeptide compared to the solubility of the first domain on its own or (iii) both to translocate the polypeptide into the cell and to increase the solubility of the polypeptide compared to the solubility of the first domain on its own, said polypeptide being free of clostridial neurotoxin and free of any clostridial neurotoxin precursor that can be converted into toxin by proteolytic action. Accordingly, the invention may thus provide a single polypeptide chain containing a domain equivalent to a clostridial toxin light chain and a domain providing the functional aspects of the H<sub>N</sub> of a clostridial toxin heavy chain, whilst lacking the functional aspects of a clostridial toxin H<sub>C</sub> domain.

For the purposes of the invention, the functional property or properties of the  $H_N$  of a clostridial toxin heavy chain that are required to be exhibited by the second domain of the polypeptide of the invention are either (i) translocation of the polypeptide into a cell, or (ii) increasing solubility of the polypeptide compared to solubility of the first domain on its own or (iii) both (i) and (ii). References hereafter to a  $H_N$  domain or to the functions of a  $H_N$  domain are references to this property or properties. The second domain is not required to exhibit other properties of the  $H_N$  domain of a clostridial toxin heavy chain.

A polypeptide of the invention can thus be soluble but lack the translocation function of a native toxin-this is of use in providing an immunogen for vaccinating or assisting to vaccinate an individual against challenge by toxin. In a specific embodiment of the invention described in an example below a polypeptide designated LH<sub>423</sub>/A elicited neutralising antibodies against type A neurotoxin. A polypeptide of the invention can likewise thus be relatively insoluble but retain the translocation function of a native toxin - this is of use if solubility is imparted to a composition made up of that polypeptide and one or more other components by one or more of said other components.

The first domain of the polypeptide of the invention cleaves one or more vesicle or plasma-membrane associated proteins essential to the specific cellular process of exocytosis, and cleavage of these proteins results in inhibition of exocytosis, typically in a non-cytotoxic manner. The cell or cells affected are not restricted to a particular type or subgroup but can include both neuronal and non-neuronal cells. The activity of clostridial neurotoxins in inhibiting exocytosis has, indeed, been observed almost universally in eukaryotic cells expressing a relevant cell surface receptor, including such diverse cells as from Aplysia (sea slug), Drosophila (fruit fly) and mammalian nerve cells, and the activity of the first domain is to be understood as including a corresponding range of cells.

The polypeptide of the invention may be obtained by expression of a recombinant nucleic acid, preferably a DNA, and is a single polypeptide, that is to say not

cleaved into separate light and heavy chain domains. The polypeptide is thus available in convenient and large quantities using recombinant techniques.

In a polypeptide according to the invention, said first domain preferably comprises a clostridial toxin light chain or a fragment or variant of a clostridial toxin light chain. The fragment is optionally an N-terminal, or C-terminal fragment of the light chain, or is an internal fragment, so long as it substantially retains the ability to cleave the vesicle or plasma-membrane associated protein essential to exocytosis. The minimal domains necessary for the activity of the light chain of clostridial toxins are described in J. Biol. Chem., Vol.267, No. 21, July 1992, pages 14721-14729. The variant has a different peptide sequence from the light chain or from the fragment, though it too is capable of cleaving the vesicle or plasma-membrane associated protein. It is conveniently obtained by insertion, deletion and/or substitution of a light chain or fragment thereof. In embodiments of the invention described below a variant sequence comprises (i) an N-terminal extension to a clostridial toxin light chain or fragment (ii) a clostridial toxin light chain or fragment modified by alteration of at least one amino acid (iii) a C-terminal extension to a clostridial toxin light chain or fragment, or (iv) combinations of 2 or more of (i)-(iii).

In further embodiments of the invention, the variant contains an amino acid sequence modified so that (a) there is no protease sensitive region between the LC and H<sub>N</sub> components of the polypeptide, or (b) the protease sensitive region is specific for a particular protease. This latter embodiment is of use if it is desired to activate the endopeptidase activity of the light chain in a particular environment or cell. Though, in general, the polypeptides of the invention are activated prior to administration.

The first domain preferably exhibits endopeptidase activity specific for a substrate selected from one or more of SNAP-25, synaptobrevin/VAMP and syntaxin. The clostridial toxin is preferably botulinum toxin or tetanus toxin.

In an embodiment of the invention described in an example below, the toxin light

chain and the portion of the toxin heavy chain are of botulinum toxin type A. In a further embodiment of the invention described in an example below, the t xin light chain and the portion of the toxin heavy chain are of botulinum toxin type B. The polypeptide optionally comprises a light chain or fragment or variant of one toxin type and a heavy chain or fragment or variant of another toxin type.

In a polypeptide according to the invention said second domain preferably comprises a clostridial toxin heavy chain  $H_N$  portion or a fragment or variant of a clostridial toxin heavy chain  $H_N$  portion. The fragment is optionally an N-terminal or C-terminal or internal fragment, so long as it retains the function of the  $H_N$  domain. Teachings of regions within the  $H_N$  responsible for its function are provided for example in Biochemistry 1995, 34, pages 15175-15181 and Eur. J. Biochem, 1989, 185, pages 197-203. The variant has a different sequence from the  $H_N$  domain or fragment, though it too retains the function of the  $H_N$  domain. It is conveniently obtained by insertion, deletion and/or substitution of a  $H_N$  domain or fragment thereof. In embodiments of the invention, described below, it comprises (i) an N-terminal extension to a  $H_N$  domain or fragment, (iii) a C-terminal extension to a  $H_N$  domain or fragment by alteration of at least one amino acid, or (iv) combinations of 2 or more of (i)-(iii). The clostridial toxin is preferably botulinum toxin or tetanus toxin.

The invention also provides a polypeptide comprising a clostridial neurotoxin light chain and a N-terminal fragment of a clostridial neurotoxin heavy chain, the fragment preferably comprising at least 423 of the N-terminal amino acids of the heavy chain of botulinum toxin type A, 417 of the N-terminal amino acids of the heavy chain of botulinum toxin type B or the equivalent number of N-terminal amino acids of the heavy chain of other types of clostridial toxin such that the fragment possesses an equivalent alignment of homologous amino acid residues.

These polypeptides of the invention are thus not composed of two or more polypeptides, linked for example by di-sulphide bridges into composite molecules. Instead, these polypeptides are single chains and are not active or their activity is

significantly reduced in an in vitro assay of neurotoxin endopeptidase activity.

Further, the polypeptides may be susceptible to be converted into a form exhibiting endopeptidase activity by the action of a proteolytic agent, such as trypsin. In this way it is possible to control the endopeptidase activity of the toxin light chain.

In a specific embodiment of the invention described in an example below, there is provided a polypeptide lacking a portion designated  $H_{\rm C}$  of a clostridial toxin heavy chain. This portion, seen in the naturally produced toxin, is responsible for binding of toxin to cell surface receptors prior to internalisation of the toxin. This specific embodiment is therefore adapted so that it can not be converted into active toxin, for example by the action of a proteolytic enzyme. The invention thus also provides a polypeptide comprising a clostridial toxin light chain and a fragment of a clostridial toxin heavy chain, said fragment being not capable of binding to those cell surface receptors involved in the intoxicating action of clostridial toxin, and it is preferred that such a polypeptide lacks an intact portion designated  $H_{\rm C}$  of a clostridial toxin heavy chain.

In further embodiments of the invention there are provided compositions containing a polypeptide comprising a clostridial toxin light chain and a portion designated  $H_N$  of a clostridial toxin heavy chain, and wherein the composition is free of clostridial toxin and free of any clostridial toxin precursor that may be converted into clostridial toxin by the action of a proteolytic enzyme. Examples of these compositions include those containing toxin light chain and  $H_N$  sequences of botulinum toxin types A, B, C<sub>1</sub>, D, E, F and G.

The polypeptides of the invention are conveniently adapted to bind to, or include, a ligand for targeting to desired cells. The polypeptide optionally comprises a sequence that binds to, for example, an immunoglobulin. A suitable sequence is a tandem repeat synthetic IgG binding domain derived from domain B of Staphylococcal protein A. Choice of immunoglobulin specificity then determines the target for a polypeptide - immunoglobulin complex. Alternatively, the

polypeptide comprises a non-clostridial sequence that binds to a cell surface receptor, suitable sequences including insulin-like growth factor-1 (IGF-1) which binds to its specific receptor on particular cell types and the 14 amino acid residue sequence from the carboxy-terminus of cholera toxin A subunit which is able to bind the cholera toxin B subunit and thence to GM1 gangliosides. A polypeptide according to the invention thus, optionally, further comprises a third domain adapted for binding of the polypeptide to a cell.

In a second aspect the invention provides a fusion protein comprising a fusion of (a) a polypeptide of the invention as described above with (b) a second polypeptide adapted for binding to a chromatography matrix so as to enable purification of the fusion protein using said chromatography matrix. It is convenient for the second polypeptide to be adapted to bind to an affinity matrix, such as a glutathione Sepharose, enabling rapid separation and purification of the fusion protein from an impure source, such as a cell extract or supernatant.

One possible second purification polypeptide is glutathione-S-transferase (GST), and others will be apparent to a person of skill in the art, being chosen so as to enable purification on a chromatography column according to conventional techniques.

As noted above, by proteolytic treatment, for example using trypsin, of a polypeptide of the invention it is possible to induce endopeptidase activity in the treated polypeptide. A third aspect of the invention provides a composition comprising a derivative of a clostridial toxin, said derivative retaining at least 10% of the endopeptidase activity of the clostridial toxin, said derivative further being non-toxic *in vivo* due to its inability to bind to cell surface receptors, and wherein the composition is free of any component, such as toxin or a further toxin derivative, that is toxic *in vivo*. The activity of the derivative preferably approaches that of natural toxin, and is thus preferably at least 30% and most preferably at least 60% of natural toxin. The overall endopeptidase activity of the composition will, of course, also be determined by the amount of the derivative that is present.

While it is known to treat naturally produced clostridial toxin to remove the  $H_{\rm C}$  domain, this treatment does not totally remove toxicity of the preparation, instead some residual toxin activity remains. Natural toxin treated in this way is therefore still not entirely safe. The composition of the invention, derived by treatment of a pure source of polypeptide advantageously is free of toxicity, and can conveniently be used as a positive control in a toxin assay, as a vaccine against clostridial toxin or for other purposes where it is essential that there is no residual toxicity in the composition.

The invention enables production of the polypeptides and fusion proteins of the invention by recombinant means.

A fourth aspect of the invention provides a nucleic acid encoding a polypeptide or a fusion protein according to any of the aspects of the invention described abov.

In one embodiment of this aspect of the invention, a DNA sequence provided to code for the polypeptide or fusion protein is not derived from native clostridial sequences, but is an artificially derived sequence not preexisting in nature.

A specific DNA (SEQ ID NO: 1) described in more detail below encodes a polypeptide or a fusion protein comprising nucleotides encoding residues 1-871 of a botulinum toxin type A. Said polypeptide comprises the light chain domain and the first 423 amino acid residues of the amino terminal portion of a botulinum toxin type A heavy chain. This recombinant product is designated LH<sub>423</sub>/A (SEQ ID NO: 2).

In a second embodiment of this aspect of the invention a DNA sequence which codes for the polypeptide or fusion protein is derived from native clostridial sequences but codes for a polypeptide or fusion protein not found in nature.

A specific DNA (SEQ ID NO: 19) described in more detail below encodes a polypeptide r a fusion protein and comprises nucleotides encoding residues 1-

1171 of a botulinum toxin type B. Said polypeptide c mprises the light chain domain and the first 728 amino acid residues of the amino terminal protein of a botulinum type B heavy chain. This recombinant product is designated LH<sub>728</sub>/B (SEQ ID NO: 20).

The invention thus also provides a method of manufacture of a polypeptide comprising expressing in a host cell a DNA according to the third aspect of the invention. The host cell is suitably not able to cleave a polypeptide or fusion protein of the invention so as to separate light and heavy toxin chains; for example, a non-clostridial host.

The invention further provides a method of manufacture of a polypeptide comprising expressing in a host cell a DNA encoding a fusion protein as described above, purifying the fusion protein by elution through a chromatography column adapted to retain the fusion protein, eluting through said chromatography column a ligand adapted to displace the fusion protein and recovering the fusion protein. Production of substantially pure fusion protein is thus made possible. Likewise, the fusion protein is readily cleaved to yield a polypeptide of the invention, again in substantially pure form, as the second polypeptide may conveniently be removed using the same type of chromatography column.

The LH<sub>N</sub>/A derived from dichain native toxin requires extended digestion with trypsin to remove the C-terminal 1/2 of the heavy chain, the  $H_{\rm C}$  domain. The loss of this domain effectively renders the toxin inactive *in vivo* by preventing its interaction with host target cells. There is, however, a residual toxic activity which may indicate a contaminating, trypsin insensitive, form of the whole type A neurotoxin.

In contrast, the recombinant preparations of the invention are the product of a discreet, defined gene coding sequence and can not be contaminated by full length toxin protein. Furthermore, the product as recovered from *E. coli*, and from other recombinant expression hosts, is an inactive single chain peptide or if expression

hosts produce a processed, active polypeptide it is not a toxin. Endopeptidase activity of LH<sub>423</sub>/A, as assessed by the current *in vitro* peptide cleavage assay, is wholly dependent on activation of the recombinant molecule between residues 430 and 454 by trypsin. Other proteolytic enzymes that cleave between these two residues are generally also suitable for activation of the recombinant molecule. Trypsin cleaves the peptide bond C-terminal to Arginine or C-terminal to Lysine and is suitable as these residues are found in the 430-454 region and are exposed (see Fig. 12).

The recombinant polypeptides of the invention are potential therapeutic agents for targeting to cells expressing the relevant substrate but which are not implicated in effecting botulism. An example might be where secretion of neurotransmitter is inappropriate or undesirable or alternatively where a neuronal cell is hyperactive in terms of regulated secretion of substances other than neurotransmitter. In such an example the function of the H<sub>c</sub> domain of the native toxin could be replaced by an alternative targeting sequence providing, for example, a cell receptor ligand and/or translocation domain.

One application of the recombinant polypeptides of the invention will be as a reagent component for synthesis of therapeutic molecules, such as disclosed in WO-A-94/21300. The recombinant product will also find application as a non-toxic standard for the assessment and development of *in vitro* assays for detection of functional botulinum or tetanus neurotoxins either in foodstuffs or in environmental samples, for example as disclosed in EP-A-0763131.

A further option is addition, to the C-terminal end of a polypeptide of the invention, of a peptide sequence which allows specific chemical conjugation to targeting ligands of both protein and non-protein origin.

In yet a further embodiment an alternative targeting ligand is added to the N-terminus of polypeptides of the invention. Recombinant  $LH_N$  derivatives have been designated that have specific protease cleavage sites engineered at the C-terminus

of the LC at the putative trypsin sensitive region and also at the extr me C-terminus of the complete protein product. These sites will enhance the activational specificity of the recombinant product such that the dichain species can only be activated by proteolytic cleavage of a more predictable nature than use of trypsin.

The LH<sub>N</sub> enzymatically produced from native BoNT/A is an efficient immunogen and thus the recombinant form with its total divorce from any full length neurotoxin represents a vaccine component. The recombinant product may serve as a basal reagent for creating defined protein modifications in support of any of the above areas.

Recombinant constructs are assigned distinguishing names on the basis of their amino acid sequence length and their Light Chain (L-chain, L) and Heavy Chain (H-chain, H) content as these relate to translated DNA sequences in the public domain or specifically to SEQ ID NO: 2 and SEQ ID NO: 20. The 'LH' designation is followed by '/X' where 'X' denotes the corresponding clostridial toxin serotype or class, e.g. 'A' for botulinum neurotoxin type A or 'TeTx' for tetanus toxin. Sequence variants from that of the native toxin polypeptide are given in parenthesis in standard format, namely the residue position number prefixed by the residue of the native sequence and suffixed by the residue of the variant.

Subscript number prefixes indicate an amino-terminal (N-terminal) extension, or where negative a deletion, to the translated sequence. Similarly, subscript number suffixes indicate a carboxy terminal (C-terminal) extension or where negative numbers are used, a deletion. Specific sequence inserts such as protease cleavage sites are indicated using abbreviations, e.g. Factor Xa is abbreviated to FXa. L-chain C-terminal suffixes and H-chain N-terminal prefixes are separated by a / to indicate the predicted junction between the L and H-chains. Abbreviations for engineered ligand sequences are prefixed or suffixed to the clostridial L-chain or H-chain corresponding to their position in the translation product.

Following this nomenclature,

- LH<sub>423</sub>/A = SEQ ID NO: 2, containing the entire L-chain and 423 amino acids of the H-chain of botulinum neurotoxin type A;
  - <sub>2</sub>LH<sub>423</sub>/A = a variant of this molecule, containing a two amino acid extension to the N-terminus of the L-chain;
  - <sub>2</sub>L<sub>/2</sub>H<sub>423</sub>/A = a further variant in which the molecule contains a two amino acid extension on the N-terminus of both the L-chain and the H-chain;
- a further variant containing a two amino acid extension to the N-terminus of the L-chain, and a Factor Xa cleavage sequence at the C-terminus of the L-chain which, after cleavage of the molecule with Factor Xa leaves a two amino acid N-terminal extension to the H-chain component; and
- $_2L_{FXe/2}H_{423}/A$ -IGF-1 = a variant of this molecule which has a further C-terminal extension to the H-chain, in this example the insulin-like growth factor 1 (IGF-1) sequence.

There now follows description of specific embodiments of the invention, illustrated by drawings in which:

- Fig. 1 shows a schematic representation of the domain structure of botulinum neurotoxin type A (BoNT/A);
- Fig. 2 shows a schematic representation of assembly of the gene for an embodiment of the invention designated LH<sub>423</sub>/A;

- Fig. 3 is a graph comparing activity of native toxin, trypsin g nerated "native"  $LH_N/A$  and an embodiment of the invention designated  ${}_2LH_{423}/A$  ( $Q_2E,N_{26}K,A_{27}Y$ ) in an *in vitro* peptide cleavage assay;
- Fig. 4 is a comparison of the first 33 amino acids in published sequences of native toxin and embodiments of the invention;
- Fig. 5 shows the transition region of an embodiment of the invention designated L/4H423/A illustrating insertion of four amino acids at the N-terminus of the H<sub>N</sub> sequence; amino acids coded for by the *Eco* 47 III restriction endonuclease cleavage site are marked and the H<sub>N</sub> sequence then begins ALN...;
- Fig. 6 shows the transition region of an embodiment of the invention designated  $L_{FXa/3}H_{423}/A$  illustrating insertion of a Factor Xa cleavage site at the C-terminus of the L-chain, and three additional amino acids coded for at the N-terminus of the H-sequence; the N-terminal amino acid of the cleavage-activated  $H_N$  will be cysteine;
- Fig. 7 shows the C-terminal portion of the amino acid sequence of an embodiment of the invention designated  $L_{FX=1/3}H_{423}/A$ -IGF-1, a fusion protein; the IGF-1 sequence begins at position  $G_{882}$ ;
- Fig. 8 shows the C-terminal portion of the amino acid sequence of an embodiment of the invention designated  $L_{FXa/3}H_{423}/A$ -CtxA14, a fusion protein; the C-terminal CtxA sequence begins at position  $Q_{882}$ ;
- Fig.9 shows the C-terminal portion of the amino acid sequence of an

embodiment of the invention designated  $L_{FXa/3}H_{423}/A$ -ZZ, a fusion protein; th C-terminal ZZ sequence begins at position  $A_{890}$  immediately after a genenase recognition site (underlined);

show schematic representations of manipulations of

Figs. 10 & 11 polypeptides of the invention; Fig. 10 shows LH<sub>423</sub>/A with N-terminal addition of an affinity purification peptide (in this case GST) and C-terminal addition of an Ig binding domain; protease cleavage sites R1, R2 and R3 enable selective enzymatic separation of domains; Fig. 11 shows specific examples of protease cleavage sites R1, R2 and R3 and a C-terminal fusion peptide

Fig. 12 shows the trypsin sensitive activation region of a polypeptide of the invention;

sequence;

shows Western blot analysis of recombinant LH<sub>107</sub>/B expressed from *E.coli*; panel A was probed with anti-BoNT/B antiserum; Lane 1, molecular weight standards; lanes 2 & 3, native BoNT/B; lane 4, immunopurified LH<sub>107</sub>/B; panel B was probed with anti-T7 peptide tag antiserum; lane 1, molecular weight standards; lanes 2 & 3, positive control *E.coli* T7 expression; lane 4 immunopurified LH<sub>107</sub>/B.

The sequence listing that accompanies this application contains the following sequences:-

SEQ ID NO:

Sequence

1

DNA coding for LH<sub>423</sub>/A

| 2           | LH <sub>423</sub> /A   |     |
|-------------|--|-----|
| 3           | DNA coding for 23LH423/A (Q2E,N26K,A27Y), of which   | an  |
|             | N-terminal portion is shown in Fig. 4.   | an  |
| 4           | <sub>23</sub> LH <sub>423</sub> /A (Q <sub>2</sub> E,N <sub>26</sub> K,A <sub>27</sub> Y)  |     |
| 5           | DNA coding for ${}_{2}LH_{423}/A$ (Q ${}_{2}E,N_{26}K,A_{27}Y$ ), of which an              | NI. |
|             | terminal portion is shown in Fig.4   | IV- |
| 6           | <sub>2</sub> LH <sub>423</sub> /A (Q <sub>2</sub> E,N <sub>26</sub> K,A <sub>27</sub> Y)   |     |
| 0           | 2L11423/A (\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\  |     |
| X-          |  |     |
| 7           | DNA coding for native BoNT/A according to Binz et a  | اد  |
| 8           | native BoNT/A according to Binz et al  | a,  |
| 9           | DNA coding for L <sub>/4</sub> H <sub>423</sub> /A   |     |
| 10          | L <sub>/4</sub> H <sub>423</sub> /A  |     |
| 11          | DNA coding for L <sub>FXB</sub> / <sub>3</sub> H <sub>423</sub> /A                         | ٠., |
| 12          | L <sub>FXa</sub> / <sub>3</sub> H <sub>423</sub> /A  |     |
| 13          | DNA coding for L <sub>FXa</sub> / <sub>3</sub> H <sub>423</sub> /A-IGF-1                   |     |
| 14          | L <sub>FXa</sub> / <sub>3</sub> H <sub>423</sub> /A-IGF-1                                  |     |
| 15          | DNA coding for L <sub>FXa</sub> / <sub>3</sub> H <sub>423</sub> /A-CtxA14                  |     |
| 16          | L <sub>FXa</sub> / <sub>3</sub> H <sub>423</sub> /A-CtxA14                                 |     |
| 17          | DNA coding for L <sub>FXe/3</sub> H <sub>423</sub> /A-ZZ                                   |     |
| 18          | L <sub>FXp/3</sub> H <sub>423</sub> /A-ZZ  | ٠.  |
|             | -FX8/3**423***   |     |
| 19          | DNA coding for LH <sub>728</sub> /B  |     |
| 20          | LH <sub>728</sub> /B   | •   |
| 21          | DNA coding for LH <sub>417</sub> /B  |     |
| 22          | LH <sub>417</sub> /B   |     |
| 23          | DNA coding for LH <sub>107</sub> /B  |     |
| 24          | LH <sub>107</sub> /B   |     |
| 25          | DNA coding for LH <sub>423</sub> /A (Q <sub>2</sub> E,N <sub>26</sub> K,A <sub>27</sub> Y) |     |
| 26          | LH <sub>423</sub> /A (Q <sub>2</sub> E,N <sub>26</sub> K,A <sub>27</sub> Y)                |     |
| 27          | DNA coding for LH <sub>417</sub> /B wherein the first 274 bases a                          | are |
| <del></del> | 717  |     |

WO 98/07864

28

modifi d to have an *E.coli* codon bias

DNA coding for LH<sub>417</sub>/B wherein bases 691-1641 of the native BoNT/B sequence have been replaced by a degenerate DNA coding for amino acid residues 231-547 of the native BoNT/B polypeptide

### Example 1

A 2616 base pair, double stranded gene sequence (SEQ ID NO: 1) has been assembled from a combination of synthetic, chromosomal and polymerase-chain-reaction generated DNA (Figure 2). The gene codes for a polypeptide of 871 amino acid residues corresponding to the entire light-chain (LC, 448 amino acids) and 423 residues of the amino terminus of the heavy-chain ( $H_c$ ) of botulinum neurotoxin type A. This recombinant product is designated the LH<sub>423</sub>/A fragment (SEQ ID NO: 2).

### Construction of the recombinant product

The first 918 base pairs of the recombinant gene were synthesised by concatenation of short oligonucleotides to generate a coding sequence with an E. coli codon bias. Both DNA strands in this region were completely synthesised as short overlapping oligonucleotides which were phosphorylated, annealed and ligated to generate the full synthetic region ending with a unique Kpnl restriction site. The remainder of the  $LH_{423}/A$  coding sequence was PCR amplified from total chromosomal DNA from  $Clostridium\ botulinum\$ and annealed to the synthetic portion of the gene.

The internal PCR amplified product sequences were then deleted and replaced with the native, fully sequenced, regions from clones of *C. botulinum* chromosomal origin to generate the final gene construct. The final composition is synthetic DNA (bases 1-913), polymerase amplified DNA (bases 914-1138 and 1976-2616) and the remainder is of *C. botulinum* chromosomal origin (bases 1139-1975). The

assembled gene was then fully sequenced and cloned into a variety of *E.coli* plasmid vectors for expression analysis.

- 18 -

# Expression of the recombinant gene and recovery of protein product

The DNA is expressed in *E. coli* as a single nucleic acid transcript producing a soluble single chain polypeptide of 99,951 Daltons predicted molecular weight. The gene is currently expressed in *E. coli* as a fusion to the commercially available coding sequence of glutathione S-transferase (GST) of *Schistosoma japonicum* but any of an extensive range of recombinant gene expression vectors such as pEZZ18, pTrc99, pFLAG or the pMAL series may be equally effective as might expression in other prokaryotic or eukaryotic hosts such as the Gram positive bacilli, the yeast *P. pastoris* or in insect or mammalian cells under appropriate conditions.

Currently, E. coli harbouring the expression construct is grown in Luria-Bertani broth (L-broth pH 7.0, containing 10 g/l bacto-tryptone, 5 g/l bacto-yeast extract and 10 g/l sodium chloride) at 37° C until the cell density (biomass) has an optical absorbance of 0.4- 0.6 at 600 nm and the cells are in mid-logarithmic growth Expression of the gene is then induced by addition of isopropylthio-β-D-galactosidase (IPTG) to a final concentration of 0.5 mM. Recombinant gene expression is allowed to proceed for 90 min at a reduced temperature of 25°C. The cells are then harvested by centrifugation, are resuspended in a buffer solution containing 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 0.5 M NaCl, 10 mM EGTA, 0.25% Tween, pH 7.0 and then frozen at -20°C. For extraction of the recombinant protein the cells are disrupted by sonication. The cell extract is then cleared of debris by centrifugation and the cleared supernatant fluid containing soluble recombinant fusion protein (GST- LH<sub>423</sub>/A) is stored at -20°C pending purification. A proportion of recombinant material is not released by the sonication procedure and this probably reflects insolubility or inclusion body formation. Currently we do not extract this material for analysis but if desired this could be readily achieved using methods known to those skilled in the art.

The recombinant GST- LH<sub>423</sub>/A is purified by adsorption onto a commercially prepared affinity matrix of glutathione Sepharose and subsequent elution with reduced glutathione. The GST affinity purification marker is then removed by proteolytic cleavage and reabsorption to glutathione Sepharose; recombinant LH<sub>423</sub>/A is recovered in the non-adsorbed material.

### Construct variants

A variant of the molecule,  $LH_{423}/A$  ( $Q_2E,N_{26}K,A_{27}Y$ ) (SEQ ID NO: 26) has been produced in which three amino acid residues have been modified within the light chain of  $LH_{423}/A$  producing a polypeptide containing a light chain sequence different to that of the published amino acid sequence of the light chain of BoNT/A .

Two further variants of the gene sequence that have been expressed and the corresponding products purified are  $_{23}LH_{423}/A$  ( $Q_2E,N_{26}K,A_{27}Y$ ) (SEQ ID NO: 4) which has a 23 amino acid N-terminal extension as compared to the predicted native L-chain of BoNT/A and  $_2LH_{423}/A$  ( $Q_2E,N_{26}K,A_{27}Y$ ) (SEQ ID NO: 6) which has a 2 amino acid N-terminal extension (Figure 4).

In yet another variant a gene has been produced which contains a *Eco* 47 III restriction site between nucleotides 1344 and 1345 of the gene sequence given in (SEQ ID NO: 1). This modification provides a restriction site at the position in the gene representing the interface of the heavy and light chains in native neurotoxin, and provides the capability to make insertions at this point using standard restriction enzyme methodologies known to those skilled in the art. It will also be obvious to those skilled in the art that any one of a number of restriction sites could be so employed, and that the *Eco* 47 III insertion simply exemplifies this approach. Similarly, it would be obvious for one skilled in the art that insertion of a restriction site in the manner described could be performed on any gene of the invention. The gene described, when expressed, codes for a polypeptide, L<sub>/4</sub>H<sub>423</sub>/A (SEQ ID NO: 10), which contains an additional four amino acids between amino acids 448 and 449 of LH<sub>423</sub>/A at a position equivalent to the amino terminus of the

heavy chain of native BoNT/A.

A variant of the gene has been expressed, L<sub>FXa/3</sub>H<sub>423</sub>/A (SEQ ID NO: 12), in which a specific proteolytic cleavage site was incorporated at the carboxy-terminal end of the light chain domain, specifically after residue 448 of L<sub>14</sub>H<sub>423</sub>/A. The cleavage site incorporated was for Factor Xa protease and was coded for by modification of SEQ ID NO: 1. It will be apparent to one skilled in the art that a cleavage site for another specified protease could be similarly incorporated, and that any gene sequence coding for the required cleavage site could be employed. Modification of the gene sequence in this manner to code for a defined protease site could be performed on any gene of the invention.

Variants of  $L_{FXa/3}H_{423}/A$  have been constructed in which a third domain is present at the carboxy-terminal end of the polypeptide which incorporates a specific binding activity into the polypeptide.

Specific examples described are:

- (1)  $L_{FXa/3}H_{423}/A$ -IGF-1 (SEQ ID NO: 14), in which the carboxy-terminal domain has a sequence equivalent to that of insulin-like growth factor-1 (IGF-1) and is able to bind to the insulin-like growth factor receptor with high affinity;
- (2)  $L_{FXa/3}H_{423}/A$ -CtxA14 (SEQ ID NO: 16), in which the carboxy-terminal domain has a sequence equivalent to that of the 14 amino acids from the carboxy-terminus of the A-subunit of cholera toxin (CtxA) and is thereby able to interact with the cholera toxin B-subunit pentamer; and
- (3)  $L_{\rm FXa/3}H_{\rm 423}/A$ -ZZ (SEQ ID NO: 18), in which the carboxy-terminal domain is a tandem repeating synthetic IgG binding domain. This variant also exemplifies another modification applicable to the current invention, namely the inclusion in the gene of a sequence coding for a protease cleavage site located between the end of the clostridial heavy chain sequence and the sequence coding for the binding

- 21 -

ligand. Specifically in this example a sequence is inserted at nucleotides 2650 to 2666 coding for a genenase cleavage site. Expression of this gene produces a polypeptide which has the desired protease sensitivity at the interface between the domain providing  $H_{\rm N}$  function and the binding domain. Such a modification enables selective removal of the C-terminal binding domain by treatment of the polypeptide with the relevant protease.

It will be apparent that any one of a number of such binding domains could be incorporated into the polypeptide sequences of this invention and that the above examples are merely to exemplify the concept. Similarly, such binding domains can be incorporated into any of the polypeptide sequences that are the basis of this invention. Further, it should be noted that such binding domains could be incorporated at any appropriate location within the polypeptide molecules of th invention.

Further embodiments of the invention are thus illustrated by a DNA of the invention further comprising a desired restriction endonuclease site at a desired location and by a polypeptide of the invention further comprising a desired protease cleavage site at a desired location.

The restriction endonuclease site may be introduced so as to facilitate further manipulation of the DNA in manufacture of an expression vector for expressing a polypeptide of the invention; it may be introduced as a consequence of a previous step in manufacture of the DNA; it may be introduced by way of modification by insertion, substitution or deletion of a known sequence. The consequence of modification of the DNA may be that the amino acid sequence is unchanged, or may be that the amino acid sequence is changed, for example resulting in introduction of a desired protease cleavage site, either way the polypeptide retains its first and second domains having the properties required by the invention.

Figure 10 is a diagrammatic representation of an expression product exemplifying features described in this example. Specifically, it illustrates a single polypeptid

incorporating a domain equivalent to the light chain of botulinum n urotoxin type A and a domain equivalent to the  $H_N$  domain of the heavy chain of botulinum neurotoxin type A with a N-terminal extension providing an affinity purification domain, namely GST, and a C-terminal extension providing a ligand binding domain, namely an IgG binding domain. The domains of the polypeptide are spatially separated by specific protease cleavage sites enabling selective enzymatic separation of domains as exemplified in the Figure. This concept is more specifically depicted in Figure 11 where the various protease sensitivities are defined for the purpose of example.

### Assay of product activity

The LC of botulinum neurotoxin type A exerts a zinc-dependent endopeptidase activity on the synaptic vesicle associated protein SNAP-25 which it cleaves in a specific manner at a single peptide bond. The  $_2LH_{423}/A$  ( $Q_2E,N_{26}K,A_{27}Y$ ) (SEQ ID NO: 6) cleaves a synthetic SNAP-25 substrate *in vitro* under the same conditions as the native toxin (Figure 3). Thus, the modification of the polypeptide sequence of  $_2LH_{423}/A$  ( $Q_2E,N_{26}K,A_{27}Y$ ) relative to the native sequence and within the minimal functional LC domains does not prevent the functional activity of the LC domains.

This activity is dependent on proteolytic modification of the recombinant GST- $_2$ LH $_{423}$ /A ( $Q_2$ E, $N_{26}$ K, $A_{27}$ Y) to convert the single chain polypeptide product to a disulphide linked dichain species. This is currently done using the proteolytic enzyme trypsin. The recombinant product (100-600  $\mu$ g/ml) is incubated at 37°C for 10-50 minutes with trypsin (10  $\mu$ g/ml) in a solution containing 140 mM NaCl, 2.7 mM KCl, 10 mM Na $_2$ HPO $_4$ , 1.8 mM KH $_2$ PO $_4$ , pH 7.3. The reaction is terminated by addition of a 100-fold molar excess of trypsin inhibitor. The activation by trypsin generates a disulphide linked dichain species as determined by polyacrylamide gel electrophoresis and immunoblotting analysis using polyclonal anti-botulinum neurotoxin type A antiserum.

<sub>2</sub>LH<sub>423</sub>/A is mor stable in the pres nce of trypsin and more activ in the in vitro

WO 98/07864 PCT/GB97/02273 - 23 -

peptide cleavage assay than is 23LH423/A. Both variants, however, are fully functional in the *in vitro* peptide cleavage assay. This demonstrates that the recombinant molecule will tolerate N-terminal amino acid extensions and this may be expanded to other chemical or organic moieties as would be obvious to those skilled in the art.

### Example 2

As a further exemplification of this invention a number of gene sequences have been assembled coding for polypeptides corresponding to the entire light-chain and varying numbers of residues from the amino terminal end of the heavy chain of botulinum neurotoxin type B. In this exemplification of the disclosure the gene sequences assembled were obtained from a combination of chromosomal and polymerase-chain-reaction generated DNA, and therefore have the nucleotide sequence of the equivalent regions of the natural genes, thus exemplifying the principle that the substance of this disclosure can be based upon natural as well as a synthetic gene sequences.

The gene sequences relating to this example were all assembled and expressed using methodologies as detailed in Sambrook J, Fritsch E F & Maniatis T (1989) Molecular Cloning: A Laboratory Manual (2nd Edition), Ford N, Nolan C, Ferguson M & Ockler M (eds), Cold Spring Harbor Laboratory Press, New York, and known to those skilled in the art.

A\_gene has been assembled coding for a polypeptide of 1171 amino acids corresponding to the entire light-chain (443 amino acids) and 728 residues from the amino terminus of the heavy chain of neurotoxin type B. Expression of this gene produces a polypeptide, LH<sub>728</sub>/B (SEQ ID NO: 20), which lacks the specific neuronal binding activity of full length BoNT/B.

A gene has also been assembled coding for a variant polypeptide,  $LH_{417}/B$  (SEQ ID NO: 22), which possesses an amino acid sequence at its carboxy terminus

equivalent by amino acid homology to that at the carboxy-terminus of the heavy chain fragment in native  $\text{LH}_\text{N}/\text{A}$  .

A gene has also been assembled coding for a variant polypeptide,  $LH_{107}/B$  (SEQ ID NO: 24), which expresses at its carboxy-terminus a short sequence from the amino terminus of the heavy chain of BoNT/B sufficient to maintain solubility of the expressed polypeptide.

### **Construct Variants**

A variant of the coding sequence for the first 274 bases of the gene shown in SEQ ID NO: 21 has been produced which whilst being a non-native nucleotide sequence still codes for the native polypeptide.

Two double stranded, a 268 base pair and a 951 base pair, gene sequences have been created using an overlapping primer PCR strategy. The nucleotide bias of these sequences was designed to have an *E.coli* codon usage bias.

For the first sequence, six oligonucleotides representing the first (5') 268 nucleotides of the native sequence for botulinum toxin type B were synthesised. For the second sequence 23 oligonucleotides representing internal sequence nucleotides 691-1641 of the native sequence for botulinum toxin type B were synthesised. The oligonucleotides ranged from 57-73 nucleotides in length. Overlapping regions, 17-20 nucleotides, were designed to give melting temperatures in the range 52-56°C. In addition, terminal restriction endonuclease sites of the synthetic products were constructed to facilitate insertion of these products into the exact corresponding region of the native sequence. The 268 bp 5' synthetic sequence has been incorporated into the gene shown in SEQ ID NO: 21 in place of the original first 268 bases (and is shown in SEQ ID NO: 27). Similarly the sequence could be inserted into other genes of the examples.

Another variant sequence equivalent to nucleotides 691 to 1641 of SEQ ID NO: 21

, and employing non-native codon usage whilst coding for a native polypeptide sequence, has been constructed using the internal synth tic sequence. This sequence (SEQ ID NO: 28) can be incorporated, alone or in combination with other variant sequences, in place of the equivalent coding sequence in any of the genes of the example.

### Example 3

An exemplification of the utility of this invention is as a non-toxic and effective immunogen. The non-toxic nature of the recombinant, single chain material was demonstrated by intraperitoneal administration in mice of GST-2LH423/A. The polypeptide was prepared and purified as described above. The amount of immunoreactive material in the final preparation was determined by enzyme linked immunosorbent assay (ELISA) using a monoclonal antibody (BA11) reactive against a conformation dependent epitope on the native LH<sub>N</sub>/A. The recombinant material was serially diluted in phosphate buffered saline (PBS; NaCl 8 g/l, KCl 0.2 g/l, Na<sub>2</sub>HPO<sub>4</sub> 1.15 g/l, KH<sub>2</sub>PO<sub>4</sub> 0.2 g/l, pH 7.4) and 0.5 ml volumes injected into 3 groups of 4 mice such that each group of mice received 10, 5 and 1 micrograms of material respectively. Mice were observed for 4 days and no deaths were seen.

For immunisation, 20  $\mu$ g of GST-<sub>2</sub>LH<sub>423</sub>/A in a 1.0 ml volume of water-in-oil emulsion (1:1 vol:vol) using Freund's complete (primary injections only) or Freund's incomplete adjuvant was administered into guinea pigs via two sub-cutaneous dorsal injections. Three injections at 10 day intervals were given (day 1, day 10 and day 20) and antiserum collected on day 30. The antisera were shown by ELISA to be immunoreactive against native botulinum neurotoxin type A and to its derivative LH<sub>N</sub>/A. Antisera which were botulinum neurotoxin reactive at a dilution of 1:2000 were used for evaluation of neutralising efficacy in mice. For neutralisation assays 0.1 ml of antiserum was diluted into 2.5 ml of gelatine phosphate buffer (GPB; Na<sub>2</sub>HPO<sub>4</sub> anhydrous 10 g/l, gelatin (Difco) 2 g/l, pH 6.5-6.6) containing a dilution range from 0.5  $\mu$ g (5X10-6 g) to 5 picograms (5X10-12 g). Aliquots of 0.5 ml w re injected into mice intraperitoneally and deaths recorded

over a 4 day period. The results are shown in Table 1 and Table 2. It can clearly be seen that 0.5 ml of 1:40 diluted anti-  $GST_{-2}LH_{423}/A$  antiserum can protect mice against intraperitoneal challenge with botulinum neurotoxin in the range 5 pg - 50 ng (1 - 10,000 mouse LD50; 1 mouse LD50 = 5 pg).

TABLE 1. Neutralisation of botulinum neurotoxin in mice by guinea pig anti-GST-2LH423/A antiserum.

# Botulinum Toxin/mouse Survivors 0.5μg 0.005μg 0.5ng 0.005ng 5pg Control (no toxin) 1 0 4 4 4 4 4 4 2 4 4 4 4 4 4 3 4 4 4 4 4 4

TABLE 2. Neutralisation of botulinum neurotoxin in mice by non-immune guinea pig antiserum.

|                     |       | <u> </u> | Sotulinum Tox | <u>in/mouse</u> |             |            |                    |
|---------------------|-------|----------|---------------|-----------------|-------------|------------|--------------------|
| Survivors<br>On Day | 0.5µg | 0.005µg  | 0.0005µg      | 0.5ng           | 0.005ng     | 5pg        | Control (no toxin) |
| 1                   | 0     | O        | 0             | 0               | o           | 2          | 4                  |
| 2                   | •     | -        | •             | •               | -           | 0          | 4                  |
| 3                   | •     | <b>-</b> | •             | -               | •           | •          | 4                  |
| . 4                 |       | •        | •             | •               | n. <b>-</b> | . <b>-</b> | 4                  |

### Example 4

Expression of recombinant LH<sub>107</sub>/B in E. coli.

As an exemplification of the expression of a nucleic acid coding for a  $LH_N$  of a clostridial neurotoxin of a serotype other than botulinum neurotoxin type A, the nucleic acid sequence (SEQ ID NO: 23) coding for the polypeptide  $LH_{107}/B$  (SEQ ID

NO: 24) was inserted into the commercially available plasmid pET28a (N vogen, Madison, WI, USA). The nucleic acid was expressed in *E. coli* BL21 (DE3) (New England BioLabs, Beverley, MA, USA) as a fusion protein with a N-terminal T7 fusion peptide, under IPTG induction at 1 mM for 90 minutes at 37°C. Cultures were harvested and recombinant protein extracted as described previously for LH<sub>423</sub>/A.

Recombinant protein was recovered and purified from bacterial paste lysates by immunoaffinity adsorption to an immobilised anti-T7 peptide monoclonal antibody using a T7 tag purification kit (New England bioLabs, Beverley, MA, USA). Purified recombinant protein was analysed by gradient (4-20%) denaturing SDS-polyacrylamide gel electrophoresis (Novex, San Diego, CA, USA) and western blotting using polyclonal anti-botulinum neurotoxin type antiserum or anti-T7 antiserum. Western blotting reagents were from Novex, immunostained proteins were visualised using the Enhanced Chemi-Luminescence system (ECL) from Amersham. The expression of an anti-T7 antibody and anti-botulinum neurotoxin type B antiserum reactive recombinant product is demonstrated in Figure 13.

The recombinant product was soluble and retained that part of the light chain responsible for endopeptidase activity.

The invention thus provides recombinant polypeptides useful inter alia as immunogens, enzyme standards and components for synthesis of molecules as described in WO-A-94/21300.

### SEQUENCE LISTING

### (1) GENERAL INFORMATION:

- (i) APPLICANT:
  - (A) NAME: MICROBIOLOGICAL RESEARCH AUTHORITY
  - (B) STREET: Centre For Applied Microbiology And Research, Porton Down
  - (C) CITY: Salisbury
  - (D) STATE: Wiltshire
  - (E) COUNTRY: UK
  - (F) POSTAL CODE (ZIP): SP4 0JG
  - (A) NAME: THE SPEYWOOD LABORATORY LIMITED
  - (B) STREET: 14 Kensington Square
  - (C) CITY: London
  - (E) COUNTRY: UK
  - (F) POSTAL CODE (ZIP): W8 5HH
  - (A) NAME: FOSTER; Keith Alan
  - (B) STREET: Centre For Applied Microbiology And Research, Porton Down
  - (C) CITY: Salisbury
  - (D) STATE: Wiltshire
  - (E) COUNTRY: UK
  - (F) POSTAL CODE (ZIP): SP4 0JG
  - (A) NAME: QUINN; Conrad Padraig
  - (B) STREET: Centre For Applied Microbiology And Research, Porton Down
  - (C) CITY: Salisbury
  - (D) STATE: Wiltshire
  - (E) COUNTRY: UK
  - (F) POSTAL CODE (ZIP): SP4 0JG
  - (A) NAME: SHONE; Clifford Charles
  - (B) STREET: Centre For Applied Microbiology And Research, Porton Down
  - (C) CITY: Salisbury
  - (D) STATE: Wiltshire
  - (E) COUNTRY: UK
  - (F) POSTAL CODE (ZIP): SP4 OJG
- (ii) TITLE OF INVENTION: Recombinant Toxin Fragments
- (iii) NUMBER OF SEQUENCES: 28
  - (iv) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: Floppy disk

    - (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

### (2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2616 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2616

| (xi) | SEQUENCE | DESCRIPTION: | SEQ | ID | NO: | 1: |
|------|----------|--------------|-----|----|-----|----|
|------|----------|--------------|-----|----|-----|----|

| ATG<br>Met        | CAG<br>Gln        | TTC<br>Phe        | GTG<br>Val        | AAC<br>Asn<br>5   | AAG<br>Lys        | CAG<br>Gln        | TTC<br>Phe        | AAC<br>Asn        | TAT<br>Tyr<br>10  | AAG<br>Lys        | GAC<br>Asp        | CCT<br>Pro        | GTA<br>Val        | AAC<br>Asn<br>15  | GGT.<br>Gly       |                | 48  |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|----------------|-----|
| GTT<br>Val        | GAC<br>.Asp       | ATT<br>Ile        | GCC<br>Ala<br>20  | TAC<br>Tyr        | ATC<br>Ile        | AAA<br>Lys        | ATT<br>Ile        | CCA<br>Pro<br>25  | AAC<br>Asn        | GCC<br>Ala        | GGC<br>Gly        | CAG<br>Gln        | ATG<br>Met<br>30  | CAG<br>Gln        | CCG<br>Pro        |                | 96  |
| GTG<br>Val        | AAG<br>Lys        | GCT<br>Ala<br>35  | Phe               | AAG<br>Lys        | ATT<br>Ile        | CAT<br>His        | AAC<br>Asn<br>40  | AAA<br>Lys        | ATC<br>Ile        | TGG<br>Trp        | GTT<br>Val        | ATT<br>Ile<br>45  | CCG<br>Pro        | GAA<br>Glu        | CGC<br>Arg        |                | 144 |
| GAT<br>Asp        | ACA<br>Thr        | TTT<br>Phe        | ACG<br>Thr        | AAC<br>Asn        | CCG<br>Pro        | GAA<br>Glu<br>55  | GAA<br>Glu        | GGA<br>Gly        | GAC<br>Asp        | TTG<br>Leu        | AAC<br>Asn<br>60  | CCG<br>Pro        | CCG<br>Pro        | CCG<br>Pro        | GAA<br>Glu        | *              | 192 |
| GCA<br>Ala<br>65  | AAG<br>Lys        | CAG<br>Gln        | GTG<br>Val        | CCA<br>Pro        | GTT<br>Val<br>70  | TCA<br>Ser        | TAC<br>Tyr        | TAC               | GAT<br>Asp        | TCA<br>Ser<br>75  | ACC<br>Thr        | TAT<br>Tyr        | CTG<br>Leu        | AGC<br>Ser        | ACA<br>Thr<br>80  |                | 240 |
| GAC<br>Asp        | AAC<br>Asn        | GAG<br>Glu        | AAG<br>Lys        | GAT<br>Asp<br>85  | AAC<br>Asn        | TAC<br>Tyr        | CTG<br>Leu        | AAG<br>Lys        | GGA<br>Gly<br>90  | GTG<br>Val        | ACC<br>Thr        | AAA<br>Lys        | TTA<br>Leu        | TTC<br>Phe<br>95  | GAG<br>Glu        | <del>.</del> . | 288 |
| CGT<br>Arg        | ATT               | TAT<br>Tyr        | TCC<br>Ser<br>100 | ACT<br>Thr        | GAC<br>Asp        | CTG<br>Leu        | GGC<br>Gly        | CGT<br>Arg<br>105 | ATG<br>Met        | CTG<br>Leu        | CTG<br>Leu        | ACC<br>Thr        | TCA<br>Ser<br>110 | ATC<br>Ile        | GTC<br>Val        | : · .          | 336 |
| CGC<br>Arg        | GGA<br>Gly        | ATC<br>Ile<br>115 | Pro               | TTT<br>Phe        | TGG<br>Trp        | GGT<br>Gly        | GGC<br>Gly<br>120 | AGT<br>Ser        | ACC<br>Thr        | ATT<br>Ile        | GAC<br>Asp        | ACG<br>Thr<br>125 | GAG<br>Glu        | TTG<br>Leu        | AAG<br>Lys        |                | 384 |
| GTT<br>Val        | ATT<br>Ile<br>130 | GAC<br>Asp        | ACT               | AAC<br>Asn        | TGC<br>Cys        | ATT<br>Ile<br>135 | AAC<br>Asn        | GTG<br>Val        | ATC<br>Ile        | CAA<br>Gln        | CCA<br>Pro<br>140 | GAC<br>Asp        | GGT<br>Gly        | AGC<br>Ser        | TAC<br>Tyr        |                | 432 |
| AGA<br>Arg<br>145 | TCT<br>Ser        | GAA<br>Glu        | GAA<br>Glu        | CTT<br>Leu        | AAC<br>Asn<br>150 | CTC<br>Leu        | GTA<br>Val        | ATC<br>Ile        | ATC<br>Ile        | GGG<br>Gly<br>155 | CCC               | TCC<br>Ser        | GCG<br>Ala        | GAC<br>Asp        | ATT<br>Ile<br>160 |                | 480 |
| ATC<br>Ile        | CAG<br>Gln        | TTT<br>Phe        | GAG<br>Glu        | TGC<br>Cys<br>165 | AAG<br>Lys        | AGC<br>Ser        | TTT               | Gly               | CAC<br>His<br>170 | GAA<br>Glu        | GTG<br>Val        | TTG<br>Leu        | AAC<br>Asn        | CTG<br>Leu<br>175 | ACG<br>Thr        |                | 528 |
| CGT               | AAC<br>Asn        | GGT<br>Gly        | TAC<br>Tyr<br>180 | GGC<br>Gly        | TCT<br>Ser        | ACT               | CAG<br>Gln        | TAC<br>Tyr<br>185 | Ile               | CGT               | TTC               | AGC<br>Ser        | CCA<br>Pro<br>190 | GAC<br>Asp        | TTC               |                | 576 |
| ACG               | TTC<br>Phe        | GGT<br>Gly<br>195 | Phe               | GAG<br>Glu        | GAG<br>Glu        | AGC<br>Ser        | CTG<br>Leu<br>200 | Glu               | GTT<br>Val        | GAT<br>Asp        | ACC<br>Thr        | AAC<br>Asn<br>205 | CCG               | CTG<br>Leu        | TTG               |                | 624 |
| GGI               | GCA<br>Ala<br>210 | Gly               | AAG<br>Lys        | TTC<br>Phe        | GCA<br>Ala        | ACT<br>Thr<br>215 | Asp               | CCA<br>Pro        | GCG<br>Ala        | GTG<br>Val        | ACC<br>Thr<br>220 | Leu               | GCA<br>Ala        | CAC               | GAG<br>Glu        |                | 672 |
| CTO<br>Lev<br>225 |                   | CAC               | GCC<br>Ala        | GGT               | CAT<br>His<br>230 | Arg               | CTG<br>Lev        | TAT<br>Tyr        | GGC<br>Gly        | Ile<br>235        | Ala               | ATT               | AAC<br>Asn        | CCG<br>Pro        | AAC<br>Asn<br>240 | .9             | 720 |

- 31 -

| AI                | g va               | I Pn               | ie ry                 | 24                  | ii As                 | n Tn:                 | r As:             | n Al                  | 25                    | r Ty:<br>0            | r Gl              | u Me               | t Se               | 25                | _                     | 1    |
|-------------------|--------------------|--------------------|-----------------------|---------------------|-----------------------|-----------------------|-------------------|-----------------------|-----------------------|-----------------------|-------------------|--------------------|--------------------|-------------------|-----------------------|------|
| GA<br>Gl          | A GT<br>u Va       | A AG<br>l Se       | C TI<br>r Ph<br>26    | ie GT               | G GA<br>u Gl          | A CTO                 | G CGG             | C ACC<br>g Th:<br>26! | r Pho                 | C GG<br>e Gly         | r GG<br>y Gl      | C CA<br>y Hi       | T GA<br>s As<br>27 | p Al              | G AAG<br>.a Lys       | 816  |
| TT<br>Phe         | r AT               | C GA<br>e As<br>27 | p se                  | C TI<br>r Le        | G CA                  | G GAC                 | AA(<br>AS1<br>28( | n Glu                 | TTO<br>Phe            | c CG                  | r CT              | G TA<br>u Ty<br>28 | r Ty               | C TA              | C AAC<br>T Asn        | 864  |
| AAC<br>Lys        | 7T':<br>Phe<br>290 | а тА               | A GA<br>s As          | T AT<br>p Il        | T GCZ<br>e Ala        | A AGT<br>a Ser<br>295 | Thi               | CTC<br>Leu            | AAC<br>Asr            | AAC<br>Lys            | GC:<br>Ala<br>300 | a Ly:              | G TC               | C'AT              | T GTG<br>e Val        | 912  |
| GGT<br>Gly<br>305 | Ini                | C AC               | T GC                  | T TC.<br>a Se       | A TTA<br>r Leu<br>310 | i Gin                 | TAT               | ATG<br>Met            | Lys                   | AAT<br>Asn<br>315     | va]               | TTT:               | r AA;              | A GA<br>s Gl      | G AAA<br>u Lys<br>320 | 960  |
| TAT<br>Tyr        | Lev                | CTI<br>Let         | A TC                  | r GA<br>r Glu<br>32 | ı Asl                 | ACA<br>Thr            | TCT<br>Ser        | GGA<br>Gly            | AAA<br>Lys<br>330     | Phe                   | TCC<br>Ser        | GTA<br>Val         | A GA:              | T AA<br>D Lys     | A TTA<br>s Leu        | 1008 |
| <b>AAA</b><br>Lys | TTT<br>Phe         | GAT<br>Asp         | C AA(<br>5 Lys<br>34( | s Lei               | TAC<br>Tyr            | Lys                   | ATG<br>Met        | TTA<br>Leu<br>345     | ACA<br>Thr            | GAG<br>Glu            | ATT               | TAC<br>Tyr         | ACA<br>Thr         | Gli               | G GAT                 | 1056 |
| AAT<br>Asn        | TTT<br>Phe         | GTT<br>Val<br>355  | . Буб                 | TTI<br>Phe          | TTT<br>Phe            | AAA<br>Lys            | GTA<br>Val<br>360 | CTT<br>Leu            | AAC<br>Asn            | AGA<br>Arg            | AAA<br>Lys        | ACA<br>Thr<br>365  | Tyr                | Lev               | AAT<br>ASD            | 1104 |
| TTT<br>Phe        | GAT<br>Asp<br>370  | гуs                | GCC<br>Ala            | GTA<br>Val          | TTT<br>Phe            | AAG<br>Lys<br>375     | ATA<br>Ile        | AAT<br>Asn            | ATA<br>Ile            | GTA<br>Val            | CCT<br>Pro<br>380 | AAG<br>Lys         | GTA<br>Val         | AA1<br>Asn        | TAC                   | 1152 |
| ACA<br>Thr<br>385 | ATA<br>Ile         | TAT                | GAT<br>Asp            | GGA<br>Gly          | TTT<br>Phe<br>390     | AAT<br>Asn            | TTA<br>Leu        | AGA<br>Arg            | AAT<br>Asn            | ACA<br>Thr<br>395     | AAT<br>Asn        | TTA<br>Leu         | GCA<br>Ala         | GCA<br>Ala        | AAC<br>Asn<br>400     | 1200 |
| TTT<br>Phe        | AAT<br>Asn         | GGT<br>Gly         | CAA<br>Gln            | AAT<br>Asn<br>405   | Thr                   | GAA<br>Glu            | ATT<br>Ile        | AAT<br>Asn            | AAT<br>Asn<br>410     | ATG<br>Met            | AAT<br>Asn        | TTT<br>Phe         | ACT<br>Thr         | AAA<br>Lys<br>415 | CTA<br>Leu            | 1248 |
| AAA<br>Lys        | TAA<br>Asn         | TTT<br>Phe         | ACT<br>Thr<br>420     | GGA<br>Gly          | TTG<br>Leu            | TTT<br>Phe            | GAA<br>Glu        | TTT<br>Phe<br>425     | TAT<br>Tyr            | AAG<br>Lys            | TTG<br>Leu        | CTA<br>Leu         | TGT<br>Cys<br>430  | GTA<br>Val        | AGA<br>Arg            | 1296 |
| GGG<br>Gly        | ATA<br>Ile         | ATA<br>Ile<br>435  | ACT<br>Thr            | TCT<br>Ser          | AAA<br>Lys            | ACT<br>Thr            | AAA<br>Lys<br>440 | TCA<br>Ser            | TTA<br>Leu            | GAT<br><b>A</b> sp    | AAA<br>Lys        | GGA<br>Gly<br>445  | TAC<br>Tyr         | AAT<br>Asn        | AAG<br>Lys            | 1344 |
| Ala               | TTA<br>Leu<br>450  | AAT<br>Asn         | GAT<br>Asp            | TTA<br>Leu          | TGT<br>Cys            | ATC<br>Ile<br>455     | AAA<br>Lys        | GTT<br>Val            | AAT<br>Asn            | Asn                   | TGG<br>Trp<br>460 | GAC<br>Asp         | TTG<br>Leu         | TTT<br>Phe        | TTT<br>Phe            | 1392 |
| AGT<br>Ser<br>465 | CCT<br>Pro         | TCA<br>Ser         | GAA<br>Glu            | GAT<br>Asp          | AAT<br>Asn<br>470     | TTT .<br>Phe '        | ACT<br>Thr        | AAT<br>Asn            | qaA                   | CTA .<br>Leu .<br>475 | AAT<br>Asn        | AAA<br>Lys         | GGA<br>Gly         | GAA<br>Glu        | GAA<br>Glu<br>480     | 1440 |
| ATT I             | ACA<br>Thr         | TCT<br>Ser         | GAT<br>Asp            | ACT<br>Thr<br>485   | AAT<br>Asn            | ATA (                 | GAA<br>Glu        | Ala .                 | GCA (<br>Ala (<br>490 | GAA (<br>Glu (        | GAA<br>Glu        | AAT<br>Asn         | Ile                | AGT<br>Ser<br>495 | TTA<br>Leu            | 1488 |
| GAT 1             | TTA<br>Leu         | ATA<br>Ile         | CAA<br>Gln<br>500     | CAA<br>Gln          | TAT<br>Tyr            | TAT ?<br>Tyr 1        | Leu '             | ACC Thr 1             | Phe                   | AAT :<br>Asn 1        | Phe               | Asp .              | AAT<br>Asn<br>510  | GAA<br>Glu        | CCT<br>Pro            | 1536 |

| GAA<br>Glu        | AAT<br>Asn        | ATT<br>Ile<br>515 | TCA<br>Ser        | ATA<br>Ile        | GAA<br>Glu        | AAT<br>Asn        | CTT<br>Leu<br>520 | TCA<br>Ser        | AGT<br>Ser        | GAC<br>Asp        | ATT<br>Ile        | ATA<br>Ile<br>525 | GGC<br>Gly        | CAA<br>Gln        | TTA<br>Leu        |     | 1584  |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-----|-------|
| GAA<br>Glu        | CTT<br>Leu<br>530 | ATG<br>Met        | CCT               | AAT<br>Asn        | ATA<br>Ile        | GAA<br>Glu<br>535 | AGA<br>Arg        | TTT<br>Phe        | CCT               | AAT<br>Asn        | GGA<br>Gly<br>540 | AAA<br>Lys        | AAG<br>Lys        | TAT<br>Tyr        | GAG<br>Glu        | • . | 1632  |
| TTA<br>Leu<br>545 | GAT<br>Asp        | AAA<br>Lys        | TAT<br>Tyr        | ACT<br>Thr        | ATG<br>Met<br>550 | TTC<br>Phe        | CAT<br>His        | TAT<br>Tyr        | CTT<br>Leu        | CGT<br>Arg<br>555 | GCT<br>Ala        | CAA<br>Gln        | GAA<br>Glu        | TTT<br>Phe        | GAA<br>Glu<br>560 |     | 1680  |
| CAT<br>His        | GGT<br>Gly        | AAA<br>Lys        | TCT<br>Ser        | AGG<br>Arg<br>565 | ATT<br>Ile        | GCT<br>Ala        | TTA<br>Leu        | ACA<br>Thr        | AAT<br>Asn<br>570 | TCT<br>Ser        | GTT<br>Val        | AAC<br>Asn        | GAA<br>Glu        | GCA<br>Ala<br>575 | TTA<br>Leu        |     | 1728  |
| TTA<br>Leu        | AAT<br>Asn        | CCT<br>Pro-       | AGT<br>Ser<br>580 | CGT<br>Arg        | GTT<br>Val        | TAT<br>Tyr        | ACA<br>Thr        | TTT<br>Phe<br>585 | TTT<br>Phe        | TCT<br>Ser        | TCA<br>Ser        | GAC<br>Asp        | TAT<br>Tyr<br>590 | GTA<br>Val        | AAG<br>Lys        |     | 1776  |
| AAA<br>Lys        | GTT<br>Val        | AAT<br>Asn<br>595 | AAA<br>Lys        | GCT<br>Ala        | ACG<br>Thr        | GAG<br>Glu        | GCA<br>Ala<br>600 | GCT<br>Ala        | ATG<br>Met        | TTT<br>Phe        | TTA<br>Leu        | GGC<br>Gly<br>605 | TGG<br>Trp        | GTA<br>Val        | GAA<br>Glu        |     | 1824  |
| CAA<br>Gln        | TTA<br>Leu<br>610 | GTA<br>Val        | TAT<br>Tyr        | GAT<br>Asp        | TTT<br>Phe        | ACC<br>Thr<br>615 | GAT<br>Asp        | GAA<br>Glu        | ACT<br>Thr        | AGC<br>Ser        | GAA<br>Glu<br>620 | GTA<br>Val        | AGT<br>Ser        | ACT<br>Thr        | ACG .             |     | 1872  |
| GAT<br>Asp<br>625 | AAA<br>Lys        | ATT<br>Ile        | GCG<br>Ala        | GAT<br>Asp        | ATA<br>Ile<br>630 | ACT<br>Thr        | ATA<br>Ile        | ATT<br>Ile        | ATT<br>Ile        | CCA<br>Pro<br>635 | TAT               | ATA<br>Ile        | GGA<br>Gly        | CCT<br>Pro        | GCT<br>Ala<br>640 |     | 1920  |
| TTA<br>Leu        | AAT<br>Asn        | ATA<br>Ile        | GGT<br>Gly        | AAT<br>Asn<br>645 | ATG<br>Met        | TTA<br>Leu        | TAT<br>Tyr        | AAA<br>Lys        | GAT<br>Asp<br>650 | GAT<br>Asp        | TTT<br>Phe        | GTA<br>Val        | GGT<br>Gly        | GCT<br>Ala<br>655 | TTA               |     | 1968  |
| Ile               | Phe               | Ser               | Gly<br>660        | Ala               | Val               | Ile               | Leu               | Leu<br>665        | Glu               | Phe               | Ile               | Pro               | Glu<br>670        | Tle               |                   |     | 2016  |
| Ile               | Pro               | Val<br>675        | Leu               | Gly               |                   | Phe               | Ala<br>680        | Leu               | Val               | Ser               | Tyr               | Ile<br>685        | Ala               | Asn               | Lys               |     | 2064  |
| GTT<br>Val        | CTA<br>Leu<br>690 | ACC               | GTT<br>Val        | CAA<br>Gln        | ACA<br>Thr        | ATA<br>Ile<br>695 | GAT<br>Asp        | AAT<br>Asn        | GCT<br>Ala        | TTA               | AGT<br>Ser<br>700 | AAA<br>Lys        | AGA<br>Arg        | AAT<br>Asn        | GAA<br>Glu        |     | 2112. |
| AAA<br>Lys<br>705 | TGG<br>Trp        | GAT<br>Asp        | GAG<br>Glu        | GTC<br>Val        | TAT<br>Tyr<br>710 | AAA<br>Lys        | TAT<br>Tyr        | ATA<br>Ile        | GTA<br>Val        | ACA<br>Thr<br>715 | AAT<br>Asn        | TGG               | TTA<br>Leu        | GCA               | AAG<br>Lys<br>720 | )   | 2160  |
| Val               | Asn               | Thr               | Gln               | 11e<br>725        | Asp               | Leu               | Ile               | Arg               | Lys<br>730        | Lys               | Met               | Lys               | Glu               | Ala<br>735        |                   |     | 2208  |
| GAA<br>Glu        | AAT<br>Asn        | CAA<br>Gln        | GCA<br>Ala<br>740 | GAA<br>Glu        | GCA<br>Ala        | ACA<br>Thr        | AAG<br>Lys        | GCT<br>Ala<br>745 | ATA               | ATA<br>Ile        | AAC               | TAT               | CAG<br>Gln<br>750 | TAT               | AAT<br>Asn        |     | 2256  |
| CAA<br>Gln        | TAT<br>Tyr        | ACT<br>Thr<br>755 | Glu               | GAA<br>Glu        | GAG<br>Glu        | AAA<br>Lys        | AAT<br>Asn<br>760 | Asn               | ATT               | AAT<br>Asn        | TTT<br>Phe        | AAT<br>Asn<br>765 | Ile               | GAT<br>Asp        | GAT<br>Asp        |     | 2304  |
| TTA<br>Leu        | AGT<br>Ser<br>770 | Ser               | Lys               | CTT<br>Leu        | TAA<br>neA        | GAG<br>Glu<br>775 | Ser               | ATA<br>Ile        | AAT<br>Asn        | AAA<br>Lys        | GCT<br>Ala<br>780 | Met               | ATT               | AAT<br>Asn        | ATA<br>Ile        |     | 2352  |

- 33 -

| AAT<br>Asn<br>785 | цуз               | TTT<br>Phe        | TTG<br>Leu        | AAT<br>Asn        | CAA<br>Gln<br>790 | Cys               | TCT<br>Ser        | GTT<br>Val        | TCA<br>Ser        | TAT<br>Tyr<br>795 | Leu               | ATG<br>Met        | AAT<br>Asn        | TCT<br>Ser        | ATG<br>Met<br>800 | 240          |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------|
| ATC<br>Ile        | CCT<br>Pro        | TAT               | GGT<br>Gly        | GTT<br>Val<br>805 | AAA<br>Lys        | CGG<br>Arg        | TTA<br>Leu        | GAA<br>Glu        | GAT<br>Asp<br>810 | TTT<br>Phe        | GAT<br>Asp        | GCT<br>Ala        | AGT<br>Ser        | CTT<br>Leu<br>815 | Lys               | <b>244</b> 1 |
| GAT<br>Asp        | GCA<br>Ala        | TTA<br>Leu        | TTA<br>Leu<br>820 | AAG<br>Lys        | TAT<br>Tyr        | ATA<br>Ile        | TAT<br>Tyr        | GAT<br>Asp<br>825 | AAT<br>Asn        | AGA<br>Arg        | GGA<br>Gly        | ACT<br>Thr        | TTA<br>Leu<br>830 | ATT<br>Ile        | GGT<br>Gly        | 2496         |
| CAA<br>Gln        | GTA<br>Val        | GAT<br>Asp<br>835 | AGA<br>Arg        | TTA<br>Leu        | AAA<br>Lys        | GAT<br>Asp        | AAA<br>Lys<br>840 | GTT<br>Val        | AAT<br>Asn        | AAT<br>Asn        | ACA<br>Thr        | CTT<br>Leu<br>845 | AGT<br>Ser        | ACA<br>Thr        | GAT<br>Asp        | 2544         |
| ATA<br>Ile        | CCT<br>Pro<br>850 | TTT<br>Phe        | CAG<br>Gln        | CTT<br>Leu        | TCC<br>Ser        | AAA<br>Lys<br>855 | TAC<br>Tyr        | GTA<br>Val        | GAT<br>Asp        | AAT<br>Asn        | CAA<br>Gln<br>860 | AGA<br>Arg        | TTA<br>Leu        | TTA<br>Leu        | TCT<br>Ser        | 2592         |
|                   |                   |                   | GAA<br>Glu        |                   |                   |                   | TAA<br>* .        |                   |                   |                   |                   |                   |                   |                   |                   | 2616         |

### (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 872 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly

Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro

Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg

Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu

Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr

Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu

Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val 100

Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys

Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr

Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile

Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr 165

Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe 185 Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu 200 Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu 250 Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Asn 280 Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val 295 Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys 310 Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr 380 Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn 395 Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu 410 405 Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg 425 Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro 505 Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu 520

- 35 -

Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu
530 540

Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Cla Clau Phe Ca

Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu 545 550 550 560

His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu 565 570 575

Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys 580 590

Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu
595 600 605

Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr 610 615 620

Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala 625 635 640

Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu 645 650 655

Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala 660 665 670

Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys 675 680 685

Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu 690 700

Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys 705 710 715 720

Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu 725 730 735

Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn 740 745 750

Gln Tyr Thr Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp 755 760 765

Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile 770 780

Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met 785 790 795 800

Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys 805 810 815

Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly 820 825 830

Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp 835 840 845

Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser 850 855 860

Thr Phe Thr Glu Tyr Ile Lys 865 870

(2) INFORMATION FOR SEQ ID NO: 3:

| (i) | SEQUENCE | CHARACTERISTICS: |
|-----|----------|------------------|
|-----|----------|------------------|

- (A) LENGTH: 2685 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: double
  (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:

  - (A) NAME/KEY: CDS (B) LOCATION:1..2685

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

|            | (XI)       | 250               | OFIAC             |            |            |            |            |                   |            | • •        |            |                |                   |            | •   |     | •    |
|------------|------------|-------------------|-------------------|------------|------------|------------|------------|-------------------|------------|------------|------------|----------------|-------------------|------------|-----|-----|------|
| Gly        | Ser        | Pro               | GGA<br>Gly        | Ile<br>5   | His        | Met        | Thr        | Ser               | Thr<br>10  | Arg        | Leu        | Gln            | Lys               | Leu<br>15  | Leu |     | 48   |
| Glu        | Phe        | Glu               | CTC<br>Leu<br>20  | Pro        | Gly        | Thr        | Met        | Glu<br>25         | Phe        | Val        | Asn        | Lys            | Gln<br>30         | Phe        | Asn | · . | 96   |
| Tyr        | Lys        | Asp<br>35         | CCT               | Val        | Asn        | Gly        | Val<br>40  | Asp               | Ile        | Ala        | Tyr        | .45            | Lys               | Ile        | Pro | ÷   | 144  |
| Lys        | Tyr<br>50  | Gly               | Gln               | Met        | Gln        | Pro<br>55  | Val        | Lys               | Ala        | Phe        | Lys<br>60  | Ile            | His               | Asn        |     | :   | 192  |
| Ile<br>65  | Trp        | Val               | ATT<br>Ile        | Pro        | Glu<br>70  | Arg        | Asp        | Thr               | Phe        | Thr<br>75  | Asn        | Pro            | GIT               | GIU        | .80 | н   | 240  |
| Asp        | Leu        | Asn               | CCG<br>Pro        | Pro<br>85  | Pro        | Glu        | Ala        | rys               | Gln<br>90  | Val        | Pro        | Val            | Ser               | 1yr<br>95  | Tyr |     | 288  |
| Asp        | Ser        | Thr               | Tyr<br>100        | Leu        | Ser        | Thr        | Asp        | Asn<br>105        | ·Glu       | Lys        | Asp        | Asn            | 110               | Leu        |     |     | 336  |
| Gly        | Val        | Thr<br>115        | AAA<br>Lys        | Leu        | Phe        | Glu        | 120        | Ile               | Tyr        | ser        | Thr        | 125            | Leu               | GIY        | Arg |     | 384. |
| Met        | Leu<br>130 | Leu               | ACC<br>Thr        | Ser        | Ile        | Val<br>135 | Arg        | GIA               | lie        | hio        | 140        | 1.rp           | GIY               | GIA        | Sei |     | 432  |
| Thr<br>145 | Ile        | Asp               | Thr               | Glu        | Leu<br>150 | Lys        | Val        | Ile               | Asp        | 155        | Asn        | Cys            | ire               | ASII       | 100 |     | 480  |
| Ile        | Gln        | Pro               | Asp               | Gly<br>165 | Ser        | Tyr        | Arg        | Ser               | 170        | GIU        | Leu        | . Asn          | Leu               | 175        |     |     | 528  |
| ATC        | GGG        | CCC               | TCC<br>Ser<br>180 | Ala        | GAC<br>Asp | ATT        | ATC        | CAG<br>Gln<br>185 | Pne        | GAG<br>Glu | TGC        | : AAG<br>: Lys | AGC<br>Ser<br>190 | Pne        | GGC |     | 576  |
| CAC<br>His | GAA<br>Glu | GTG<br>Val<br>195 | . Leu             | AAC<br>Asn | CTG<br>Leu | ACG<br>Thr | Arg<br>200 | Asn               | GGT<br>Gly | TAC        | GGC<br>Gly | Ser<br>205     | Thi               | CAG<br>Gln | TAC |     | 624  |

| AT                | T CG<br>e Ar<br>21 | g Pn              | C AG<br>e Se      | C CC<br>r Pr          | A GA                | C TTO<br>P Phe<br>215 | Thi                   | TTO Pho           | C GG<br>e Gl        | T TTO                 | C GAG<br>e Gl     | u Gl              | G AG<br>u Se      | C CI               | G GAC                 | G 672 |
|-------------------|--------------------|-------------------|-------------------|-----------------------|---------------------|-----------------------|-----------------------|-------------------|---------------------|-----------------------|-------------------|-------------------|-------------------|--------------------|-----------------------|-------|
| GT<br>Va:<br>22:  | I AS               | T AC<br>p Th      | C AA<br>r As      | C CC                  | G CTO<br>Lev<br>230 | т тел                 | G GG7<br>I Gly        | GC/<br>Ala        | A GGG               | C AA(<br>y Lys<br>239 | s Phe             | C GC<br>e Al      | A AC<br>a Th      | T GA<br>r As       | T CCA<br>P Pro<br>240 | )     |
| GC0<br>Ala        | G GTO              | G ACC             | C CTO             | G GCA<br>L Ala<br>245 | a His               | GAG<br>Glu            | CTO<br>Leu            | ATC<br>Ile        | C CAC<br>His<br>250 | s Ala                 | GG:               | CA'               | CG<br>Ar          | T CT<br>g Le<br>25 | G TAT<br>u Tyr<br>5   | 768   |
| GTZ               | / 116              | Ala               | 260<br>260        | e Asi                 | 1 Pro               | ) Asn                 | Arg                   | Val<br>265        | . Phe               | : Lys                 | Val               | . Ası             | 270               | r Ası              | C GCC                 |       |
| TAC               | TAC                | GA(<br>Glu<br>275 | ı Met             | AGT<br>Ser            | GGT<br>Gly          | TTA<br>Leu            | GAA<br>Glu<br>280     | GTA<br>Val        | AGC<br>Ser          | TTC<br>Phe            | GAG<br>Glu        | GAZ<br>Glu<br>285 | Let               | G CGG              | ACG<br>Thr            | 864   |
| TTO               | GGT<br>Gly<br>290  | , GT?             | CAT<br>His        | GAI<br>Asp            | GCG<br>Ala          | Lys<br>295            | TTT                   | ATC               | GAC<br>Asp          | AGC<br>Ser            | TTG<br>Leu<br>300 | Gln               | GAC<br>Glu        | AA(<br>Asi         | GAG<br>Glu            | 912   |
| Phe               | Arg                | CTG<br>Leu        | TAC<br>Tyr        | TAC<br>Tyr            | TAC<br>Tyr<br>310   | AAC<br>Asn            | AAG<br>Lys            | TTT<br>Phe        | AAA<br>Lys          | GAT<br>Asp<br>315     | ATT<br>Ile        | GCA<br>Ala        | AGT<br>Ser        | ACA<br>Thr         | CTG<br>Leu<br>320     | 960   |
| Asn               | Lys                | Ala               | : Lys             | Ser<br>325            | Ile                 | Val                   | Gly                   | Thr               | Thr<br>330          | Ala                   | Ser               | Leu               | Gln               | Tyr<br>335         |                       | 1008  |
| Lys               | Asn                | Val               | Phe 340           | Lys                   | Glu                 | Lys                   | Tyr                   | Leu<br>345        | Leu                 | Ser                   | Glu               | Asp               | Thr<br>350        | Ser                | -                     | 1056  |
| Lys               | Phe                | Ser<br>355        | Val               | Asp                   | Lys                 | TTA<br>Leu            | Lys<br>360            | Phe               | Asp                 | Lys                   | Leu               | Tyr<br>365        | Lys               | Met                | Leu                   | 1104  |
| Thr               | 370                | Ile               | Tyr               | Thr                   | Glu                 | GAT<br>Asp<br>375     | Asn                   | Phe               | Val                 | Lys                   | Phe<br>380        | Phe               | Lys               | Val                | Leu                   | 1152  |
| AAC<br>Asn<br>385 | AGA<br>Arg         | AAA<br>Lys        | ACA<br>Thr        | TAT<br>Tyr            | TTG<br>Leu<br>390   | AAT<br>Asn            | TTT<br>Phe            | GAT<br>Asp        | AAA<br>Lys          | GCC<br>Ala<br>395     | GTA<br>Val        | TTT<br>Phe        | AAG<br>Lys        | ATA<br>Ile         | AAT<br>Asn<br>400     | 1200  |
| Ile               | Val                | Pro               | Lys               | Val<br>405            | Asn                 | TAC<br>Tyr            | Thr                   | Ile               | Tyr<br>410          | <b>Asp</b>            | Gly               | Phe               | Asn               | Leu<br>415         | Arg                   | 1248  |
| AAT<br>Asn        | ACA<br>Thr         | AAT<br>Asn        | TTA<br>Leu<br>420 | GCA<br>Ala            | GCA<br>Ala          | AAC<br>Asn            | Phe .                 | AAT<br>Asn<br>425 | GGT<br>Gly          | CAA<br>Gln            | AAT<br>Asn        | ACA<br>Thr        | GAA<br>Glu<br>430 | ATT<br>Ile         | TAA<br>naA            | 1296  |
| TAA<br>naA        | ATG<br>Met         | AAT<br>Asn<br>435 | TTT<br>Phe        | ACT<br>Thr            | AAA<br>Lys          | CTA .<br>Leu :        | AAA .<br>Lys .<br>440 | AAT<br>Asn        | TTT<br>Phe          | ACT<br>Thr            | Gly               | TTG<br>Leu<br>445 | TTT<br>Phe        | GAA<br>Glu         | TTT<br>Phe            | 1344  |
| TAT               | AAG<br>Lys<br>450  | TTG<br>Leu        | CTA<br>Leu        | TGT<br>Cys            | Val .               | AGA (<br>Arg (<br>455 | GGG :                 | ATA .<br>Ile      | ATA<br>Ile          | Thr                   | TCT<br>Ser<br>460 | AAA<br>Lys        | ACT<br>Thr        | AAA<br>Lys         | TCA<br>Ser            | 1392  |
| TTA<br>Leu<br>465 | GAT<br>Asp         | AAA<br>Lys        | GGA<br>Gly        | Tyr                   | AAT<br>Asn<br>470   | AAG (<br>Lys )        | GCA :                 | ITA /             | Asn .               | GAT (<br>Asp )<br>475 | TTA (<br>Leu (    | TGT<br>Cys        | ATC<br>Ile        | Lys                | GTT<br>Val<br>480     | 1440  |

|   |                   |                   | , ma              |                   | mme               |                   |                   |                   |                   |                   |                   |                   |                   |                   | •                 |                   |     |      |
|---|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-----|------|
|   | AAT               | AA1               | Trp               | Asp               | Leu<br>485        | Pne               | Phe               | AGT<br>Ser        | Pro               | Ser<br>490        | Glu               | GAT<br>Asr        | AAT<br>Asn        | TTT<br>Phe        | T ACT             | AAT<br>Asn        |     | 1488 |
|   | GAT<br>Asp        | CTA<br>Leu        | AAT<br>Asn        | Lys<br>500        | GIY               | GAA<br>Glu        | GAA<br>Glu        | ATT               | ACA<br>Thr<br>505 | TCT<br>Ser        | GAT<br>Asp        | ACT<br>Thr        | AAT<br>Asn        | ATA<br>Ile        | : Gli             | GCA<br>Ala        |     | 1536 |
|   | GCA<br>Ala        | GAA<br>Glu        | GAA<br>Glu<br>515 | Asn               | ATT               | AGT<br>Ser        | TTA<br>Leu        | GAT<br>Asp<br>520 | Leu               | ATA<br>Ile        | CAA<br>Gln        | CAA<br>Gln        | TAT<br>Tyr<br>525 | TAT<br>Tyr        | TTA<br>Leu        | ACC<br>Thr        |     | 1584 |
|   | TTT               | AAT<br>Asn<br>530 | Pne               | GAT<br>Asp        | AAT<br>Asn        | GAA<br>Glu        | CCT<br>Pro<br>535 | GAA<br>Glu        | AAT<br>Asn        | ATT               | TCA<br>Ser        | ATA<br>Ile<br>540 | GAA<br>Glu        | AAT<br>Asn        | CTI<br>Leu        | TCA<br>Ser        | • . | 1632 |
|   | AGT<br>Ser<br>545 | GAC<br>Asp        | ATT<br>Ile        | ATA<br>Ile        | GIA               | CAA<br>Gln<br>550 | TTA<br>Leu        | GAA<br>Glu        | CTT<br>Leu        | ATG<br>Met        | CCT<br>Pro<br>555 | AAT<br>Asn        | ATA               | GAA<br>Glu        | AGA<br>Arg        | TTT<br>Phe<br>560 |     | 1680 |
|   | CCT<br>Pro        | AAT<br>Asn        | GGA<br>Gly        | AAA<br>Lys        | AAG<br>Lys<br>565 | TAT<br>Tyr        | GAG<br>Glu        | TTA<br>Leu        | GAT<br>Asp        | AAA<br>Lys<br>570 | TAT<br>Tyr        | ACT<br>Thr        | ATG<br>Met        | TTC<br>Phe        | CAT<br>His<br>575 | TAT<br>Tyr        |     | 1728 |
|   | CTT<br>Leu        | CGT               | GCT<br>Ala        | CAA<br>Gln<br>580 | GAA<br>Glu        | TTT               | GAA<br>Glu        | CAT               | GGT<br>Gly<br>585 | AAA<br>Lys        | TCT               | AGG<br>Arg        | ATT<br>Ile        | GCT<br>Ala<br>590 | TTA<br>Leu        | ACA<br>Thr        |     | 1776 |
| - | AAT<br>Asn        | TCT<br>Ser        | GTT<br>Val<br>595 | AAC<br>Asn        | GAA<br>Glu        | GCA<br>Ala        | TTA<br>Leu        | TTA<br>Leu<br>600 | AAT<br>Asn        | CCT<br>Pro        | AGT<br>Ser        | CGT<br>Arg        | GTT<br>Val<br>605 | TAT<br>Tyr        | ACA<br>Thr        | TTT               |     | 1824 |
|   | TTT<br>Phe        | TCT<br>Ser<br>610 | TCA<br>Ser        | GAC<br>Asp        | TAT<br>Tyr        | GTA<br>Val        | AAG<br>Lys<br>615 | AAA<br>Lys        | GTT<br>Val        | AAT<br>Asn        | Lys               | GCT<br>Ala<br>620 | ACG<br>Thr        | GAG<br>Glu        | GCA<br>Ala        | GCT<br>Ala        |     | 1872 |
|   | ATG<br>Met<br>625 | TTT<br>Phe        | TTA<br>Leu        | GGC               | TGG<br>Trp        | GTA<br>Val<br>630 | GAA<br>Glu        | CAA<br>Gln        | TTA<br>Leu        | GTA<br>Val        | TAT<br>Tyr<br>635 | GAT<br>Asp        | TTT<br>Phe        | ACC<br>Thr        | GAT<br>Asp        | GAA<br>Glu<br>640 |     | 1920 |
|   | ACT<br>Thr        | AGC<br>Ser        | GAA<br>Glu        | GTA<br>Val        | AGT<br>Ser<br>645 | ACT<br>Thr        | ACG<br>Thr        | GAT<br>Asp        | AAA<br>Lys        | ATT<br>Ile<br>650 | GCG<br>Ala        | GAT<br>Asp        | ATA<br>Ile        | ACT<br>Thr        | ATA<br>Ile<br>655 | ATT<br>Ile        |     | 1968 |
|   | ATT<br>Ile        | CCA<br>Pro        | TAT<br>Tyr        | ATA<br>Ile<br>660 | GGA<br>Gly        | CCT<br>Pro        | GCT<br>Ala        | TTA<br>Leu        | AAT<br>Asn<br>665 | ATA<br>Ile        | GGT<br>Gly        | AAT<br>Asn        | Met               | TTA<br>Leu<br>670 | TAT<br>Tyr        | AAA<br>Lys        |     | 2016 |
|   | GAT<br>Asp        | GAT<br>Asp        | TTT<br>Phe<br>675 | GTA<br>Val        | GGT<br>Gly        | GCT<br>Ala        | Leu               | ATA<br>Ile<br>680 | TTT<br>Phe        | TCA<br>Ser        | GGA<br>Gly        | GCT<br>Ala        | GTT<br>Val<br>685 | ATT<br>Ile        | CTG<br>Leu        | TTA<br>Leu        |     | 2064 |
|   | GLu               | TTT<br>Phe<br>690 | ATA<br>Ile        | CCA<br>Pro        | <b>GAG</b><br>Glu | Ile               | GCA<br>Ala<br>695 | ATA<br>Ile        | CCT<br>Pro        | GTA<br>Val        | TTA<br>Leu        | GGT<br>Gly<br>700 | ACT<br>Thr        | TTT<br>Phe        | GCA<br>Ala        | CTT<br>Leu        |     | 2112 |
|   | GTA<br>Val<br>705 | TCA<br>Ser        | TAT<br>Tyr        | ATT<br>Ile        | GCG<br>Ala        | AAT<br>Asn<br>710 | AAG<br>Lys        | GTT<br>Val        | CTA<br>Leu        | Thr               | GTT<br>Val<br>715 | CAA<br>Gln        | ACA<br>Thr        | ATA<br>Ile        | GAT<br>Asp        | AAT<br>Asn<br>720 |     | 2160 |
|   | GCT<br>Ala        | TTA<br>Leu        | AGT<br>Ser        | AAA<br>Lys        | AGA<br>Arg<br>725 | AAT<br>Asn        | GAA<br>Glu        | AAA<br>Lys        | Trp .             | GAT<br>Asp<br>730 | GAG<br>Glu        | GTC<br>Val        | TAT<br>Tyr        | AAA<br>Lys        | TAT<br>Tyr<br>735 | ATA<br>Ile        |     | 2208 |
|   | GTA<br>Val        | ACA<br>Thr        | AAT<br>Asn        | TGG<br>Trp<br>740 | TTA<br>Leu        | GCA .<br>Ala      | AAG<br>Lys        | Val               | AAT<br>Asn<br>745 | ACA<br>Thr        | CAG .<br>Gln      | ATT<br>Ile        | Asp               | CTA<br>Leu<br>750 | ATA<br>Ile        | AGA<br>Arg        |     | 2256 |
|   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |     |      |

- 39 -

| AAJ<br>Lys        | A AA/<br>; Lys    | A ATC<br>5 Met<br>755 | . Lys      | GAZ<br>Glu        | GCT<br>Ala        | TTA<br>Leu        | GAA<br>Glu<br>760 | LASI       | CAI               | A GCA<br>n Ala    | A .GA.            | A GCA<br>Ala<br>765 | Thr            | AAQ<br>Lys        | GCT<br>Ala        | 2304         |
|-------------------|-------------------|-----------------------|------------|-------------------|-------------------|-------------------|-------------------|------------|-------------------|-------------------|-------------------|---------------------|----------------|-------------------|-------------------|--------------|
| ATA<br>Ile        | ATA<br>Ile<br>770 | : MSII                | TAT<br>Tyr | CAG<br>Gln        | TAT<br>Tyr        | AAT<br>Asn<br>775 | CAA<br>Gln        | TAT        | ACT<br>Thr        | GAG<br>Glu        | GAA<br>Glu<br>780 | Glu                 | AAA<br>Lys     | AAT<br>Asn        | AAT<br>Asn        | <b>235</b> 2 |
| ATT<br>Ile<br>785 | M511              | TTT<br>Phe            | AAT<br>Asn | ATT               | GAT<br>Asp<br>790 | GAT<br>Asp        | TTA<br>Leu        | AGT<br>Ser | TCG<br>Ser        | AAA<br>Lys<br>795 | Leu               | AAT<br>Asn          | GAG<br>Glu     | TCT<br>Ser        | ATA<br>Ile<br>800 | 2400         |
| AAT<br>Asn        | AAA<br>Lys        | GCT<br>Ala            | ATG<br>Met | ATT<br>Ile<br>805 | AAT<br>Asn        | ATA<br>Ile        | AAT<br>Asn        | AAA<br>Lys | TTT<br>Phe<br>810 | Leu               | AAT<br>Asn        | CAA<br>Gln          | TGC<br>Cys     | TCT<br>Ser<br>815 | GTT<br>Val        | 2448         |
| 561               | IYI               | Leu                   | 820        | ASI               | ser               | Met               | IIe               | 825        | Tyr               | Gly               | Val               | AAA<br>Lys          | <b>Arg</b> 830 | Leu               | Glu               | 2496         |
| GAT<br>Asp        | TTT<br>Phe        | GAT<br>Asp<br>835     | GCT<br>Ala | AGT<br>Ser        | CTT<br>Leu        | AAA<br>Lys        | GAT<br>Asp<br>840 | GCA<br>Ala | TTA<br>Leu        | TTA<br>Leu        | AAG<br>Lys        | TAT<br>Tyr<br>845   | ATA<br>Ile     | TAT<br>Tyr        | GAT<br>Asp        | 2544         |
| AAT<br>Asn        | AGA<br>Arg<br>850 | GGA<br>Gly            | ACT<br>Thr | TTA<br>Leu        | TTE               | GGT<br>Gly<br>855 | CAA<br>Gln        | GTA<br>Val | GAT<br>Asp        | AGA<br>Arg        | TTA<br>Leu<br>860 | AAA<br>Lys          | GAT<br>Asp     | AAA<br>Lys        | GTT<br>Val        | 2592         |
| AAT<br>Asn<br>865 | AAT<br>Asn        | ACA<br>Thr            | CTT<br>Leu | AGT<br>Ser        | ACA<br>Thr<br>870 | GAT<br>Asp        | ATA<br>Ile        | CCT<br>Pro | TTT<br>Phe        | CAG<br>Gln<br>875 | CTT<br>Leu        | TCC<br>Ser          | AAA<br>Lys     | TAC<br>Tyr        | GTA<br>Val<br>880 | 2640         |
| GAT<br>Asp        | AAT<br>Asn        | CAA<br>Gln            | Arg        | TTA<br>Leu<br>885 | TTA<br>Leu        | TCT<br>Ser        | ACA<br>Thr        | TTT<br>Phe | ACT<br>Thr<br>890 | GAA<br>Glu        | TAT               | ATT .               | Lys            | TAA<br>*<br>895   |                   | 2685         |

### (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 895 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Gly Ser Pro Gly Ile His Met Thr Ser Thr Arg Leu Gln Lys Leu Leu

Glu Phe Glu Leu Pro Gly Thr Met Glu Phe Val Asn Lys Gln Phe Asn

Tyr Lys Asp Pro Val Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro

Lys Tyr Gly Gln Met Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys 50 60

Ile Trp Val Ile Pro Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly

Asp Leu Asn Pro Pro Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr

Asp Ser Thr Tyr Leu Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys 105

Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly 185 His Glu Val Leu Asn Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr 200 Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu 215 Val Asp Thr Asn Pro Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu Leu Ile His Ala Gly His Arg Leu Tyr 245 Gly Ile Ala Ile Asn Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala . 260 Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr 280 Phe Gly Gly His Asp Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu 290 295 300 Phe Arg Leu Tyr Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser

Leu Asp Lys Gly Tyr Asn Lys Ala Leu Asn Asp Leu Cys Ile Lys Val 480 Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala

500 505 Find Ash The Glu Ala

Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr 515 520 525

Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser 530 540

Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe 545 550 555 560

Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr 565 570 575

Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr 580 585 590

Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe 595 600 605

Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala 610 615 620

Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu 625 635 640

Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile 645 650 655

Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys
660 665 670

Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu 675 680 685

Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu 690 695 700

Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn 705 710 715 720

Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile
725 730 735

Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg
740 745 750

Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala 755 760 765

Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn 770 780

Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile 785 790 795 800

Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val 805 810 815 WO 98/07864 PCT/GB97/02273 - 42 -

Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu 825

Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp 840 845

Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val 850 855

Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val 875

Asp Asn Gln Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys \* 885 890 895

### (2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2622 base pairs (B) TYPE: nucleic acid

  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:

  - (A) NAME/KEY: CDS
    (B) LOCATION:1..2622

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

|   |               |            |                  | •                |                  |            |            |                  |                  |                  |            |            |                  |                  |                  |            | •   |  |
|---|---------------|------------|------------------|------------------|------------------|------------|------------|------------------|------------------|------------------|------------|------------|------------------|------------------|------------------|------------|-----|--|
|   | GA<br>ly<br>l | TCC<br>Ser | ATG<br>Met       | GAG<br>Glu       | TTC<br>Phe<br>5  | GTG<br>Val | AAC<br>Asn | AAG<br>Lys       | CAG<br>Gln       | TTC<br>Phe<br>10 | AAC<br>Asn | TAT        | AAG<br>Lys       | GAC<br>Asp       | CCT<br>Pro<br>15 | GTA<br>Val | 48  |  |
| A | AC<br>LSII    | GGT<br>Gly | GTT<br>Val       | GAC<br>Asp<br>20 | ATT<br>Ile       | GCC<br>Ala | TAC        | ATC<br>Ile       | AAA<br>Lys<br>25 | ATT              | CCA<br>Pro | AAG<br>Lys | TAC<br>Tyr       | GGC<br>Gly<br>30 | CAG<br>Gln       | ATG<br>Met | 96  |  |
| 9 | AG<br>In      | CCG<br>Pro | GTG<br>Val<br>35 | AAG<br>Lys       | GCT<br>Ala       | TTC<br>Phe | AAG<br>Lys | ATT<br>Ile<br>40 | CAT<br>His       | AAC<br>Asn       | AAA<br>Lys | ATC<br>Ile | TGG<br>Trp<br>45 | GTT<br>Val       | ATT<br>Ile       | CCG<br>Pro | 144 |  |
|   |               |            |                  |                  | TTT<br>Phe       |            |            |                  |                  |                  |            |            |                  |                  |                  |            | 192 |  |
| F |               |            |                  |                  | CAG<br>Gln       |            |            |                  |                  |                  |            |            |                  |                  |                  |            | 240 |  |
|   |               |            |                  |                  | GAG<br>Glu<br>85 |            |            |                  |                  |                  |            |            |                  |                  |                  |            | 288 |  |
|   |               |            |                  |                  | TAT<br>Tyr       |            |            |                  |                  |                  |            |            |                  |                  |                  |            | 336 |  |
|   |               |            |                  |                  | ATC<br>Ile       |            |            |                  |                  |                  |            |            |                  |                  |                  |            | 384 |  |
|   |               |            |                  |                  | GAC<br>Asp       |            |            |                  |                  |                  |            |            |                  |                  |                  |            | 432 |  |

# WO 98/07864 PCT/GB97/02273 - 43 -

| AC<br>'Se<br>14   | - Ly              | AC A              | GA T<br>rg S       | cr             | GAA<br>Glu         | GA<br>Gl:<br>15   | u re              | T AA<br>u As       | C CT                  | C GI<br>u Va       | A AT                | .e I]             | rc G<br>le G       | GG Co                             | CC :              | rcc<br>Ser       | GCG<br>Ala<br>160 |      |
|-------------------|-------------------|-------------------|--------------------|----------------|--------------------|-------------------|-------------------|--------------------|-----------------------|--------------------|---------------------|-------------------|--------------------|-----------------------------------|-------------------|------------------|-------------------|------|
| GA<br>As          | C AI              | T A               | rc c<br>le G       | TII            | TTT<br>Phe<br>165  | GI                | G TG<br>u Cy      | C AA<br>s Ly       | G AG                  | C TT<br>r Ph<br>17 | e Gl                | C CA<br>y Hi      | AC GA<br>.s G1     | NA G                              | al L              | TG<br>Leu<br>175 | AAC<br>Asn        | 528  |
| CT<br>Le          | G AC<br>u Th      | G CC              | .g A               | AC<br>sn<br>80 | GGT<br>Gly         | TAC<br>Ty:        | C GG<br>r Gl      | C TC<br>y Se       | T AC'<br>r Th:<br>18! | r Gl               | G TA<br>n Ty        | C AT              | T CO               | T T                               | ne S              | GC<br>er         | CCA<br>Pro        | 576  |
| GA<br>As          | C TT<br>p Ph      | C AC<br>e Th      | IT EI              | TC (<br>ne (   | GGT<br>Gly         | TTC<br>Phe        | GA(               | G GA<br>1 Gl<br>20 | G AGO<br>u Sei<br>0   | CTC<br>r Lei       | G GA                | G GT<br>u Va      | T GA<br>l As<br>20 | p Th                              | C A               | AC<br>sn         | CCG<br>Pro        | 624  |
| CT                | G TT<br>u Le      | u GI              | T GO<br>Ý A        | CA (           | GGC<br>Gly         | AAG<br>Lys        | Phe<br>215        | : Al               | A ACT                 | r ga:<br>: Asi     | r cci<br>p Pro      | A GCO<br>Al.      | a Va               | G AC                              | C C               | TG<br>eu         | GCA<br>Ala        | 672  |
| CAC<br>His<br>22  | 5 GI              | G CT<br>u Le      | G AT<br>u I]       | c c            | CAC                | GCC<br>Ala<br>230 | GIY               | CAT<br>His         | r CGT                 | CTC<br>Lev         | TAT<br>1 Ty1<br>235 | Gly               | C AT               | T GC<br>e Al                      | G A'<br>a I       | TT<br>le         | AAC<br>Asn<br>240 | 720  |
| CC0<br>Pro        | AA E              | C CG              | C GI<br>g Va       | T F            | TC<br>Phe          | AAG<br>Lys        | GTT<br>Val        | AA(<br>Asr         | ACC<br>Thr            | AAC<br>Asn<br>250  | ı Ala               | TAC<br>Tyr        | TA                 | C GA                              | u Me              | rg<br>et         | AGT<br>Ser        | 768  |
| GG7<br>Gly        | TTA<br>Leu        | A GA.             | A GI<br>u Va<br>26 | ב ג            | GC<br>er           | TTC<br>Phe        | GAG<br>Glu        | GAA<br>Glu         | CTG<br>Leu<br>265     | Arg                | ACG<br>Thr          | TTC               | GG<br>Gly          | r GG(<br>/ Gl <sub>2</sub><br>27( | y Hi              | T.               | GAT<br>Asp        | 816  |
| GCG<br>Ala        | AAC<br>Lys        | Phe 275           | 3 11               | C G<br>e A     | AC<br>.sp          | AGC<br>Ser        | TTG<br>Leu        | CAG<br>Gln<br>280  | GAG<br>Glu            | AAC<br>Asn         | GAG<br>Glu          | TTC               | CG1<br>Arg<br>285  | J Lev                             | TA<br>1 Ty        | ir i             | TAC<br>Tyr        | 864  |
| TAC               | AAC<br>Asn<br>290 | πλε               | TT<br>Ph           | T A<br>e L     | AA (               | GAT<br>Asp        | тте               | Ala                | AGT<br>Ser            | Thr                | Leu                 | Asn               | Lvs                | GCT<br>Ala                        | T AA<br>Ly        | .G :             | rcc<br>Ser        | 912  |
| ATT<br>Ile<br>305 | GTG<br>Val        | GGT<br>Gly        | ACO<br>Th:         | C A            | nr 1               | GCT<br>Ala<br>310 | TCA<br>Ser        | TTA<br>Leu         | CAG<br>Gln            | TAT<br>Tyr         | ATG<br>Met<br>315   | AAA<br>Lys        | AAT<br>Asn         | GTI<br>Val                        | TT<br>Ph          | e I              | AAA<br>Lys<br>120 | 960  |
| GAG<br>Glu        | AAA<br>Lys        | TAT               | CT(                | عدا ۱          | TA S<br>eu S<br>25 | rct<br>Ser        | GAA<br>Glu        | GAT<br>Asp         | ACA<br>Thr            | TCT<br>Ser<br>330  | GGA<br>Gly          | AAA<br>Lys        | TTT<br>Phe         | TCG<br>Ser                        | GT.<br>Va<br>33   | 1 7              | TAS               | 1008 |
| AAA<br>Lys        | TTA<br>Leu        | AAA<br>Lys        | Phe<br>340         | : AE           | AT A               | AAG<br>Lys        | TTA<br>Leu        | TAC<br>Tyr         | AAA<br>Lys<br>345     | ATG<br>Met         | TTA<br>Leu          | ACA<br>Thr        | GAG<br>Glu         | ATT<br>Ile<br>350                 | Ty:               | C A              | CA<br>hr          | 1056 |
| GAG<br>Glu        | GAT<br>Asp        | AAT<br>Asn<br>355 | TTT                | GI<br>Va       | TT A               | YAG<br>Lys        | TTT<br>Phe        | TTT<br>Phe<br>360  | AAA<br>Lys            | GTA<br>Val         | CTT<br>Leu          | AAC<br>Asn        | AGA<br>Arg<br>365  | AAA<br>Lys                        | AC:               | A T              | 'AT<br>'yr        | 1104 |
| TTG<br>Leu        | AAT<br>Asn<br>370 | TTT<br>Phe        | GAT<br>Asp         | Ly<br>Ly       | NA G               | lla '             | GTA<br>Val<br>375 | TTT<br>Phe         | AAG<br>Lys            | ATA<br>Ile         | AAT<br>Asn          | ATA<br>Ile<br>380 | GTA<br>Val         | CCT<br>Pro                        | AAC<br>Lys        | G<br>S V         | TA<br>al          | 1152 |
| AAT<br>Asn<br>385 | TAC<br>Tyr        | ACA<br>Thr        | ATA<br>Ile         | TA<br>Ty       | T A                | AT (<br>sp (      | GGA<br>Gly        | TTT<br>Phe         | AAT<br>Asn            | TTA<br>Leu         | AGA<br>Arg<br>395   | TAA<br>neA        | ACA<br>Thr         | AAT<br>Asn                        | TTA               | 1 A              | CA<br>la<br>00    | 1200 |
| GCA<br>Ala        | AAC<br>Asn        | TTT<br>Phe        | AAT<br>Asn         | GG<br>G1<br>40 | уG                 | AA i              | AAT<br>Asn        | ACA<br>Thr         | GAA .<br>Glu          | ATT<br>Ile<br>410  | AAT<br>Asn          | AAT<br>Asn        | ATG<br>Met         | AAT<br>Asn                        | TTT<br>Phe<br>415 | T                | CT<br>hr          | 1248 |

PCT/GB97/02273

WO 98/07864

675

- 45 -

| no.               | 69                | 0                 | T De                  | u 111             | r val                 | 695                   | Ini               | : 116             | e Asp             | Ası               | 700               | Let               | ı Sez             | r Ly:             | A AGA<br>s Arg      | 2112 |
|-------------------|-------------------|-------------------|-----------------------|-------------------|-----------------------|-----------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---------------------|------|
| 70                | 5                 |                   | s II,                 | p As              | 710                   | )<br>Y Val            | тух               | . ràs             | Tyr               | 715               | Val               | Thr               | : Asr             | Tr                | G TTA<br>Leu<br>720 | 2160 |
| ATC               | . Ly:             | s va              | I ASI                 | 72                | 5                     | . IIe                 | Asp               | Leu               | 730               | Arg               | Lys               | Lys               | Met               | 735               |                     | 2208 |
| GCT<br>Ala        | TTA<br>Lev        | A GAI             | A AA7<br>1 Asr<br>740 | ı ÇII             | A GCA<br>n Ala        | GAA<br>Glu            | GCA<br>Ala        | ACA<br>Thr<br>745 | AAG<br>Lys        | GCT<br>Ala        | ATA<br>Ile        | ATA<br>Ile        | AAC<br>Asn<br>750 | Tyr               | CAG<br>Gln          | 2256 |
| TAT               | TAA '             | CAM<br>Glr<br>755 | ııyı                  | C ACI             | GAG<br>Glu            | GAA<br>Glu            | GAG<br>Glu<br>760 | AAA<br>Lys        | AAT<br>Asn        | AAT<br>Asn        | ATT<br>Ile        | AAT<br>Asn<br>765 | TTT<br>Phe        | AAT<br>Asn        | ATT<br>Ile          | 2304 |
| GAT<br>Asp        | GAT<br>Asp<br>770 | Der               | AGT<br>Ser            | TCG<br>Ser        | Lys                   | CTT<br>Leu<br>775     | AAT<br>Asn        | GAG<br>Glu        | TCT<br>Ser        | ATA<br>Ile        | AAT<br>Asn<br>780 | AAA<br>Lys        | GCT<br>Ala        | ATG<br>Met        | ATT<br>Ile          | 2352 |
| AAT<br>Asn<br>785 | ATA<br>Ile        | AAT<br>Asn        | AAA<br>Lys            | TTT<br>Phe        | TTG<br>Leu<br>790     | AAT<br>Asn            | CAA<br>Gln        | TGC<br>Cys        | TCT<br>Ser        | GTT<br>Val<br>795 | TCA<br>Ser        | TAT<br>Tyr        | TTA<br>Leu        | ATG<br>Met        | AAT<br>Asn<br>008   | 2400 |
| TCT<br>Ser        | ATG<br>Met        | ATC               | Pro                   | TAT<br>Tyr<br>805 | GGT<br>Gly            | GTT<br>Val            | AAA<br>Lys        | CGG<br>Arg        | TTA<br>Leu<br>810 | GAA<br>Glu        | GAT<br>Asp        | TTT<br>Phe        | GAT<br>Asp        | GCT<br>Ala<br>815 | AGT<br>Ser          | 2448 |
| Deu               | Lys,              | Asp               | 820                   | reu               | TTA<br>Leu            | ràs                   | Tyr               | 825               | Tyr               | Asp               | Asn               | Arg               | Gly<br>830        | Thr               | Leu                 | 2496 |
| ATT               | GGT<br>Gly        | CAA<br>Gln<br>835 | GTA<br>Val            | GAT<br>Asp        | AGA<br>Arg            | TTA .<br>Leu :        | Lys .             | Asp               | Lys `             | Val .             | Asn .             | Asn               | ACA<br>Thr        | CTT<br>Leu        | AGT<br>Ser          | 2544 |
| ACA<br>Thr        | GAT<br>Asp<br>850 | ATA<br>Ile        | CCT<br>Pro            | TTT<br>Phe        | CAG<br>Gln            | CTT :<br>Leu :<br>855 | rcc :             | AAA<br>Lys        | TAC (             | Val .             | GAT<br>Asp<br>860 | AAT<br>Asn        | CAA .<br>Gln .    | AGA<br>Arg        | TTA<br>Leu          | 2592 |
| TTA<br>Leu<br>865 | TCT<br>Ser        | ACA<br>Thr        | TTT<br>Phe            | ACT<br>Thr        | GAA 1<br>Glu 1<br>870 | TAT A                 | ATT I             | AAG '<br>Lys      | TAA<br>*          |                   |                   |                   |                   |                   |                     | 2622 |

# (2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 874 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
- Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val 10
- Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Lys Tyr Gly Gln Met
- Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro

| Glu                             | Arg<br>50                              | Asp   | Thr                                 | Phe                                 | Thr  | Asn<br>55                                     | Pro   | Glu                                     | Glu   | Gly                                     | <b>Asp</b> 60                          | Leu                                     | Asn                                     | Pro   | Pro  |
|---------------------------------|--|---|-------------------------------------|-------------------------------------|--|---|---|---|---|---|--|---|---|---|--|
| Pro<br>65                       | Glu                                    | Ala   | Lys                                 | Gln                                 | Val<br>70  | Pro   | Val   | Ser                                     | Tyr   | Tyr<br>75                               | Asp                                    | Ser                                     | Thr                                     | Tyr   | Leu<br>80  |
| Ser                             | Thr                                    | Asp.  | Asn                                 | Glu<br>85                           | Lys  | Asp   | Asn   | Tyr                                     | Leu<br>90   | Lys                                     | Gly                                    | Val                                     | Thr                                     | Lys<br>95   | Leu  |
| Phe                             | Glu                                    | Arg   | Ile<br>100                          | Tyr                                 | Sér  | Thr   | Asp   | Leu<br>105                              | Gly   | Arg                                     | Met                                    | Leu                                     | Leu<br>110                              | Thr   | Ser  |
| Ile                             | Val                                    | Arg<br>115                                  | Gly                                 | Ile                                 | Pro  | Phe   | Trp<br>120                                    | Gly                                     | Gly   | Ser                                     | Thr                                    | Ile<br>125                              | Asp                                     | Thr   | Glu  |
| Leu                             | Lys<br>130                             | Val   | Ile                                 | Asp                                 | Thr  | Asn<br>135                                    | Cys   | Ile                                     | Asn   | Val                                     | Ile<br>140                             | Gln                                     | Pro                                     | Asp   | Gly  |
| Ser<br>145                      | Tyr                                    | Arg   | Ser                                 | Glu                                 | Glu<br>150   | Leu   | Asn   | Leu                                     | Val   | Ile<br>155                              | Ile                                    | Gly                                     | Pro                                     | Ser   | Ala<br>160   |
| Asp                             | Ile                                    | Ile   | Gln                                 | Phe<br>165                          | Glu  | Cys   | Lys   | Ser                                     | Phe<br>170  | Gly                                     | His                                    | Glu                                     | Val                                     | Leu<br>175  | Asn  |
| Leu                             | Thr                                    | Arg   | Asn<br>180                          | Gly                                 | Tyr  | Gly   | Ser   | Thr<br>105                              |   | Tyr                                     | Ile                                    | Arg                                     | Phe<br>190                              | Ser   | Pro  |
| Asp                             | Phe                                    | Thr<br>195                                  | Phe                                 | Gly                                 | Phe  | Glu   | Glu<br>200                                    |   | Leu   | Glu                                     | Val                                    | Asp<br>205                              | Thr                                     | Asn   | Pro  |
| Leu                             | Leu<br>210                             | Gly   | Ala                                 | Gly                                 | Lys  | Phe<br>215                                    | Ala   | Thr                                     | Asp   | Pro                                     | Ala<br>220                             | Val                                     | Thr                                     | Leu   | Ala  |
|                                 |  |   | •                                   |                                     |  |   |   |   |   |   |  |   |   |   |  |
| His<br>225                      | Glu                                    | Leu   | Ile                                 | His                                 | Ala<br>230   | Gly   | His   | Arg                                     | Leu   | Tyr<br>235                              | Gly                                    | Ile                                     | Ala                                     | Ile   | Asn<br>240   |
| 225                             |  |   |                                     |                                     |  | •   |   |   |   | 235                                     |  |   |   |   | 240  |
| 225<br>Pro                      | Asn                                    | Arg   | Val                                 | Phe<br>245<br>Ser                   | 230  | Val   | Asn   | Thr                                     | Asn<br>250<br>Arg   | 235<br>Ala                              | Tyr                                    | Tyr                                     | Glu                                     | Met<br>255  | 240<br>Ser   |
| 225<br>Pro                      | Asn                                    | Arg<br>Glu                                  | Val<br>Val<br>260                   | Phe<br>245<br>Ser                   | 230<br>Lys   | Val<br>Glu                                    | Asn   | Thr<br>Leu<br>265                       | Asn<br>250<br>Arg   | Ala<br>Thr                              | Tyr<br>Phe                             | Tyr<br>Gly                              | Glu<br>Gly<br>270                       | Met<br>255<br>His   | 240<br>Ser<br>Asp                                    |
| Pro<br>Gly                      | Asn<br>Leu<br>Lys                      | Arg<br>Glu<br>Phe<br>275                    | Val<br>260<br>Ile                   | Phe<br>245<br>Ser                   | 230<br>Lys<br>Phe                                    | Val<br>Glu<br>Leu                             | Asn<br>Glu<br>Gln<br>280                      | Thr<br>Leu<br>265<br>Glu                | Asn<br>250<br>Arg   | 235<br>Ala<br>Thr<br>Glu                | Tyr<br>Phe<br>Phe                      | Tyr<br>Gly<br>Arg<br>285                | Glu<br>Gly<br>270<br>Leu                | Met<br>255<br>His<br>Tyr                                    | 240<br>Ser<br>Asp<br>Tyr                             |
| 225<br>Pro<br>Gly<br>Ala<br>Tyr | Asn<br>Leu<br>Lys<br>Asn<br>290<br>Val | Arg<br>Glu<br>Phe<br>275<br>Lys             | Val<br>260<br>Ile                   | Phe<br>245<br>Ser<br>Asp            | 230<br>Lys<br>Phe-<br>Ser                            | Val<br>Glu<br>Leu<br>Ile<br>295               | Asn<br>Glu<br>Gln<br>280<br>Ala               | Thr<br>Leu<br>265<br>Glu<br>Ser         | Asn<br>250<br>Arg<br>Asn                                    | 235<br>Ala<br>Thr<br>Glu<br>Leu         | Tyr<br>Phe<br>Phe<br>Asn<br>300        | Tyr<br>Gly<br>Arg<br>285<br>Lys         | Glu<br>Gly<br>270<br>Leu<br>Ala         | Met<br>255<br>His<br>Tyr                                    | 240<br>Ser<br>Asp<br>Tyr<br>Ser                      |
| Pro Gly Ala Tyr Ile 305         | Asn<br>Leu<br>Lys<br>Asn<br>290<br>Val | Arg Glu Phe 275 Lys Gly                     | Val<br>260<br>Ile<br>Phe<br>Thr     | Phe 245 Ser Asp Lys                 | 230<br>Lys<br>Phe<br>Ser<br>Asp                      | Val<br>Glu<br>Leu<br>Ile<br>295<br>Ser        | Asn<br>Glu<br>Gln<br>280<br>Ala<br>Leu        | Thr Leu 265 Glu Ser                     | Asn<br>250<br>Arg<br>Asn<br>Thr                             | 235 Ala Thr Glu Leu Met 315             | Tyr<br>Phe<br>Phe<br>Asn<br>300<br>Lys | Tyr Gly Arg 285 Lys Asn                 | Glu<br>Gly<br>270<br>Leu<br>Ala<br>Val  | Met<br>255<br>His<br>Tyr<br>Lys                             | 240<br>Ser<br>Asp<br>Tyr<br>Ser<br>Lys<br>320        |
| Pro Gly Ala Tyr Ile 305 Glu     | Asn<br>Leu<br>Lys<br>Asn<br>290<br>Val | Arg Glu Phe 275 Lys Gly Tyr                 | Val<br>260<br>Ile<br>Phe<br>Thr     | Phe 245 Ser Asp Lys Thr Leu 325     | 230<br>Lys<br>Phe<br>Ser<br>Asp<br>Ala<br>310        | Val<br>Glu<br>Leu<br>Ile<br>295<br>Ser<br>Glu | Asn<br>Glu<br>Gln<br>280<br>Ala<br>Leu<br>Asp | Thr Leu 265 Glu Ser Gln Thr             | Asn<br>250<br>Arg<br>Asn<br>Thr<br>Tyr<br>Ser<br>330        | 235 Ala Thr Glu Leu Met 315 Gly         | Tyr<br>Phe<br>Phe<br>Asn<br>300<br>Lys | Tyr Gly Arg 285 Lys Asn —Phe            | Glu Gly 270 Leu Ala Val Ser             | Met<br>255<br>His<br>Tyr<br>Lys<br>Phe<br>Val<br>335        | 240<br>Ser<br>Asp<br>Tyr<br>Ser<br>Lys<br>320<br>Asp |
| Pro Gly Ala Tyr Ile 305 Glu Ly8 | Asn Leu Lys Asn 290 Val Lys            | Arg Glu Phe 275 Lys Gly Tyr                 | Val 260 Ile Phe Thr Leu Phe 340     | Phe 245 Ser Asp Lys Thr Leu 325 Asp | 230<br>Lys<br>Phe<br>Ser<br>Asp<br>Ala<br>310<br>Ser | Val Glu Leu Ile 295 Ser Glu Leu               | Asn Glu Gln 280 Ala Leu Asp                   | Thr Leu 265 Glu Ser Gln Thr Lys 345 Lys | Asn<br>250<br>Arg<br>Asn<br>Thr<br>Tyr<br>Ser<br>330<br>Met | 235 Ala Thr Glu Leu Met 315 Gly Leu     | Tyr Phe Phe Asn 300 Lys Lys            | Tyr Gly Arg 285 Lys Asn —Phe Glu        | Glu Gly 270 Leu Ala Val Ser Ile 350     | Met<br>255<br>His<br>Tyr<br>Lys<br>Phe<br>Val<br>335<br>Tyr | 240 Ser Asp Tyr Ser Lys 320 Asp                      |
| Pro Gly Ala Tyr Ile 305 Glu Ly8 | Asn Leu Lys Asn 290 Val Lys Leu Asp    | Arg Glu Phe 275 Lys Gly Tyr Lys Asn 355 Phe | Val 260 Ile Phe Thr Leu Phe 340 Phe | Phe 245 Ser Asp Lys Thr Leu 325 Asp | 230 Lys Phe Ser Asp Ala 310 Ser Lys                  | Val Glu Leu Ile 295 Ser Glu Leu Phe           | Asn Glu Gln 280 Ala Leu Asp Tyr Phe 360 Phe   | Thr Leu 265 Glu Ser Gln Thr Lys 345 Lys | Asn<br>250<br>Arg<br>Asn<br>Thr<br>Tyr<br>Ser<br>330<br>Met | 235 Ala Thr Glu Leu Met 315 Gly Leu Leu | Tyr Phe Phe Asn 300 Lys Lys Thr        | Tyr Gly Arg 285 Lys Asn Phe Glu Arg 365 | Glu Gly 270 Leu Ala Val Ser Ile 350 Lys | Met<br>255<br>His<br>Tyr<br>Lys<br>Phe<br>Val<br>335<br>Tyr | 240 Ser Asp Tyr Ser Lys 320 Asp Thr                  |

- 47 -

Ala Asn Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr 410 Lys Leu Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu 455 Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly 520 Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu 570 Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr 585 Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly 635 Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu 705 710 715 720 Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln 745

240

288

|          |          |          |            |            |            |              |            |            |            |            |            |            |            |            |            | •          |     |            |
|----------|----------|----------|------------|------------|------------|--------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|------------|
| Ту       | r As     | sn       | Gln<br>755 | Tyr        | Thr        | Glu          | Glu'       | Glu<br>760 | Lys        | Asn        | Asn        | Ile        | Asn<br>765 | Phe        | Asn        | Tle        |     |            |
| As       | _        | sp<br>70 | Leu        | Ser        | Ser        | Lys          | Leu<br>775 | Asn        | Glu        | Ser        | Ile        | Asn<br>780 | Lys        | Ala        | Met        | Ile        | ٠.  |            |
| As<br>78 |          | le       | Asn        | Lys        | Phe        | Leu<br>790   | Asn        | Gln        | Cys        | Ser        | Val<br>795 | Ser        | Tyr        | Leu        | Met        | Asn<br>800 |     | •          |
| Se       | r Me     | et       | Ile        | Pro        | Tyr<br>805 | Gly          | Val        | Lys        | Arg        | Leu<br>810 | Glu        | Asp        | Phe        | Asp        | Ala<br>815 | Ser        |     |            |
| Le       | u Ly     | γs       | Asp        | Ala<br>820 | Leu        | Leu          | Lys        | Tyr        | Ile<br>825 | Tyr        | Asp        | Asn        | Arg        | Gly<br>830 | Thr        | Leu        |     |            |
| 11       | .e Gi    | •        | Gln<br>835 |            | Asp        | Arg          | Leu        | Lys<br>840 | Asp        | Lys        | Val        | Asn        | Asn<br>845 | Thr        | Leu        | Ser        |     |            |
| Th       |          | sp<br>50 | Ile        | Pro        | Phe        | Gln          | Leu<br>855 | Ser        | Lys        | Tyr        | Val        | Asp<br>860 |            | Gln        | Arg        | Leu        |     |            |
| Le<br>86 |          | er       | Thr        | Phe        | Thr        | Glu<br>870   |            | Ile        | Lys        | *          |            |            |            |            |            |            |     | · · ·      |
| (2       | ) II     | NFO      | RMAI       | CION       | FOR        | SEQ          | ID N       | 10: 7      | 7:         |            |            |            |            |            |            |            |     | <i>,</i> . |
|          |          | (i)      | SEC.       | NIENC      | יד רו      | IARAC        | TER        | STIC       | 75:        |            |            |            | •          |            |            |            |     |            |
|          |          | ( 1 /    | (2         | ) LE       | ENGTH      | 1: 26        | 313 E      | oase       | pair       | cs         |            |            |            |            |            | ٠.         |     |            |
|          |          |          |            |            |            | nuc]<br>EDNE |            |            |            |            |            | -          |            |            |            |            |     |            |
|          |          |          | (E         | ) (TC      | POLC       | GY:          | line       | ar         |            |            |            |            |            |            |            |            |     | •          |
|          | <b>(</b> | ii)      | MOI        | ECUI       | E TY       | PE:          | DNA        | .(ger      | nomic      | :)         |            |            |            |            |            |            |     |            |
|          | ( :      | ix)      | FEA        | TURE       | S:         |              |            |            | -          |            |            |            | •          |            |            |            |     |            |
|          | ,        |          |            |            |            | ŒY:<br>[ON:] |            | 513        |            |            |            |            | ٠.         |            |            |            |     |            |
|          |          |          | (-         | , 20       |            | -            |            |            |            |            |            |            |            |            |            |            |     |            |
|          | ()       | xi)      | SEC        | UENC       | E DE       | ESCRI        | PTIC       | ON: S      | SEQ 1      | D NO       | ): 7:      | :          |            |            |            |            |     |            |
| ÞΤ       | e co     | CA       | TTT        | ĠTT        | AAT        | AAA          | CAA        | TTT        | AAT        | TAT        | AAA        | GAT        | CCT        | GTA        | AAT        | GGT        |     | 4.8        |
| Me       | t Pi     | ro       | Phe        | Val        | Asn<br>5   | Lys          | Gln        | Phe        | Asn        | Tyr<br>10  | Lys        | Asp        | Pro        | Val        | Asn<br>15  | Gly        |     | •          |
| GI       | T G      | AT       | ATT        | GCT        | TAT        | ATA          | AAA        | ATT        | CCA        | AAT        | GCA        | GGA<br>Gly | CAA        | ATG        | CAA        | CCA        | , - | 96         |
| Vä       | (T M:    | вÞ       | 116        | 20         | TÅT        | 116          | Lys        | +10        | 25         | AD.II      | . ALG      | GIY.       |            | 30         | Gan        | ,          |     |            |
| GI       | A A      | AA       | GCT        | TTT        | AAA        | ATT          | CAT        | AAT        | AAA        | ATA        | TGG        | GTT        | ATT        | CCA        | GAA        | AGA        |     | 144        |
| Va       | il Ly    | ys       | Ala<br>35  | Phe        | Lys        | Ile          | His        | Asn<br>40  | Lys        | ile        | Trp        | Val        | 11e<br>45  | Pro        | GIU        | Arg        |     | *.         |
| GF       | A TA     | CA       | TTT        | ACA        | AAT        | CCT          | GAA        | GAA        | GGA        | GAT        | TTA        | AAT        | CCA        | CCA        | CCA        | GAA        |     | 192        |
| Αs       | p T      | hr       | Phe        | Thr        | Asn        | Pro          | Glu<br>55  | GIU        | GIY.       | Asp        | Leu        | Asn<br>60  | PTO        | PTO        | LLO        | Glu        |     |            |

GCA AAA CAA GTT CCA GTT TCA TAT TAT GAT TCA ACA TAT TTA AGT ACA Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr 65

GAT AAT GAA AAA GAT AAT TAT TTA AAG GGA GTT ACA AAA TTA TTT GAG Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu 85

|                   |                    |                    |                     |                       |                       |                   |                       |                    |                   | •                     |                   |                       |                    |                   |                   |          |      |
|-------------------|--------------------|--------------------|---------------------|-----------------------|-----------------------|-------------------|-----------------------|--------------------|-------------------|-----------------------|-------------------|-----------------------|--------------------|-------------------|-------------------|----------|------|
| AC<br>Az          | A AI<br>g Il       | T TA<br>e Ty       | T TC<br>T Se<br>10  | r in                  | T GA<br>r As          | T CT<br>p Lei     | r GG.<br>u Gl         | A AG<br>y Ar<br>10 | g Me              | G TT(<br>t Le         | G TT.             | A AC<br>u Th          | A TC<br>r Se<br>11 | r I               | TA GT<br>le Va    | 'A<br>.1 | 336  |
| AC<br>Ar          | G GG<br>G Gl       | A AT<br>y Il<br>11 | e Pr                | A TT<br>o Ph          | T TG<br>e Tr          | G GG7<br>p Gly    | GG/<br>Gly<br>120     | y Se               | T AC              | A ATA                 | A GA              | T AC                  | r Gl               | A TI              | A AA<br>u Ly      | A        | 384  |
| GT<br>Va          | T AT<br>1 11<br>13 | е АБ               | T AC<br>p Th        | T AA'<br>r As         | T TG                  | T ATT             | : Asi                 | r GT(              | G ATA             | A CAZ<br>e Glr        | CCI<br>Pro<br>140 | Ası                   | r GG<br>p Gl       | r AG<br>Y Se      | T TA              | T<br>r   | 432  |
| AG<br>Ar<br>14    | g Se               | A GA<br>r Gl       | A GA<br>u Gl        | A CT                  | r AA:<br>u As:<br>150 | r CTA<br>n Leu    | GT#                   | A ATA              | A ATA             | A GGA<br>e Gly<br>155 | Pro               | TCA<br>Ser            | A GCT              | r ga<br>a As      | T AT:<br>p Ile    | =        | 480  |
| AT.               | A CAG              | TT<br>n Pho        | r GAJ<br>e Glu      | A TGT<br>1 Cys<br>165 | 3 Lye                 | A AGC<br>S Ser    | TTI<br>Phe            | GG#                | A CAT<br>His      | Glu                   | GT1<br>Val        | TTO<br>Lev            | AA7<br>Asr         | CT<br>Le          | u Thi             | 3<br>c   | 528  |
| CG/<br>Arg        | A AA?<br>g Asi     | r GG:<br>1 Gl;     | TAT<br>7 Ty1<br>180 | : GI                  | C TCT                 | ACT<br>Thr        | CAA<br>Gln            | TAC<br>Tyr<br>185  | Ile               | AGA<br>Arg            | TTT<br>Phe        | AGC<br>Ser            | Pro                | As                | r TTT             | <b>C</b> | 576  |
| AC:               | TTT                | GG:<br>Gly<br>195  | Pne                 | GAG<br>Glu            | GAG<br>Glu            | TCA<br>Ser        | CTT<br>Leu<br>200     | Glu                | GTT<br>Val        | GAT<br>Asp            | ACA<br>Thr        | AAT<br>Asn<br>205     | CCT                | Lei               | TTA<br>Leu        |          | 624  |
| GG7<br>Gly        | GCA<br>Ala<br>210  | GI                 | : AAA<br>Lys        | TTT<br>Phe            | GCT<br>Ala            | ACA<br>Thr<br>215 | GAT<br>Asp            | CCA<br>Pro         | GCA<br>Ala        | GTA<br>Val            | ACA<br>Thr<br>220 | TTA<br>Leu            | GCA<br>Ala         | CA7<br>His        | GAA<br>Glu        |          | 672  |
| Leu<br>225        | LILE               | CAT<br>His         | GCT<br>Ala          | GGA<br>Gly            | CAT<br>His<br>230     | AGA<br>Arg        | TTA<br>Leu            | TAT                | GGA<br>Gly        | ATA<br>Ile<br>235     | GCA<br>Ala        | ATT<br>Ile            | TAA<br>naA         | CCA               | AAT<br>Asn<br>240 |          | 720  |
| AGG<br>Arg        | GTI<br>Val         | TTT<br>Phe         | Lys                 | GTA<br>Val<br>245     | Asn                   | ACT<br>Thr        | AAT<br>Asn            | GCC<br>Ala         | TAT<br>Tyr<br>250 | TAT<br>Tyr            | GAA<br>Glu        | ATG<br>Met            | AGT<br>Ser         | GGG<br>Gly<br>255 | Leu               |          | 768  |
| GIU               | Val                | ser                | Phe<br>260          | GIu                   | GIU                   | CTT<br>Leu        | Arg                   | Thr<br>265         | Phe               | Gly                   | Gly               | His                   | Asp<br>270         | Ala               | Lys               |          | 816  |
| Phe               | ATA<br>Ile         | GAT<br>Asp<br>275  | AGT<br>Ser          | TTA<br>Leu            | CAG<br>Gln            | GAA<br>Glu        | AAC<br>Asn<br>280     | GAA<br>Glu         | TTT<br>Phe        | CGT                   | CTA<br>Leu        | TAT<br>Tyr<br>285     | TAT<br>Tyr         | TAT               | AAT<br>Asn        |          | 864  |
| Lys               | 290                | Lys                | Asp                 | He                    | Ala                   | AGT<br>Ser<br>295 | Thr                   | Leu                | Asn               | Lys                   | Ala<br>300        | Lys                   | Ser                | Ile               | Val               |          | 912  |
| GGT<br>Gly<br>305 | ACT<br>Thr         | ACT                | GCT<br>Ala          | TCA<br>Ser            | TTA<br>Leu<br>310     | CAG<br>Gln        | TAT<br>Tyr            | ATG<br>Met         | AAA<br>Lys        | AAT<br>Asn<br>315     | GTT<br>Val        | TTT<br>Phe            | AAA<br>Lys         | GAG<br>Glu        | AAA<br>Lys<br>320 |          | 960  |
| TAT<br>Tyr        | CTC<br>Leu         | CTA<br>Leu         | TCT                 | GAA<br>Glu<br>325     | GAT<br>Asp            | ACA<br>Thr        | TCT<br>Ser            | GGA<br>Gly         | AAA<br>Lys<br>330 | TTT Phe               | TCG<br>Ser        | GTA<br>Val            | Asp                | AAA<br>Lys<br>335 | TTA<br>Leu        |          | 1008 |
| AAA<br>Lys        | TTT<br>Phe         | GAT<br>Asp         | AAG<br>Lys<br>340   | TTA<br>Leu            | TAC<br>Tyr            | AAA .<br>Lys !    | Met .                 | TTA<br>Leu<br>345  | ACA<br>Thr        | GAG :                 | ATT<br>Ile        | Tyr '                 | ACA<br>Thr<br>350  | GAG<br>Glu        | GAT<br>Asp        |          | 1056 |
| AAT<br>Asn        | TTT<br>Phe         | GTT<br>Val<br>355  | AAG<br>Lys          | TTT<br>Phe            | TTT<br>Phe            | AAA (<br>Lys :    | GTA (<br>Val :<br>360 | CTT .<br>Leu .     | AAC .<br>Asn .    | AGA 1<br>Arg 1        | Lys '             | ACA '<br>Thr '<br>365 | TAT (              | TTG<br>Leu        | AAT<br>Asn        |          | 1104 |

| Phe               | GAT<br>Asp<br>370 | AAA<br>Lys        | GCC               | GTA<br>Val        | TTT               | AAG<br>Lys<br>375 | ATA<br>Ile        | AAT<br>Asn        | ATA<br>Ile        | GTA<br>Val        | CCT<br>Pro<br>380 | AAG<br>Lys        | GTA<br>Val        | AAT<br>Asn        | TAC               |   | ļ152 |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---|------|
| ACA<br>Thr<br>385 | ATA<br>Ile        | TAT               | GAT<br>Asp        | GGA<br>Gly        | TTT<br>Phe<br>390 | AAT<br>Asn        | TTA<br>Leu        | AGA<br>Arg        | AAT<br>Asn        | ACA<br>Thr<br>395 | AAT<br>Asn        | TTA<br>Leu        | GCA<br>Ala        | GCA<br>Ala        | AAC<br>Asn<br>400 |   | 1200 |
| TTT<br>Phe        | AAT<br>Asn        | GGT<br>Gly        | CAA<br>Gln        | AAT<br>Asn<br>405 | ACA<br>Thr        | GAA<br>Glu        | ATT<br>Ile        | AAT<br>Asn        | AAT<br>Asn<br>410 | ATG<br>Met        | AAT<br>Asn        | TTT<br>Phe        | ACT<br>Thr        | AAA<br>Lys<br>415 | CTA<br>Leu        |   | 1248 |
| AAA<br>Lys        | AAT<br>Asn        | TTT<br>Phe        | ACT<br>Thr<br>420 | GGA<br>Gly        | TTG               | TTT<br>Phe        | GAA<br>Glu        | TTT<br>Phe<br>425 | TAT<br>Tyr        | AAG<br>Lys        | TTG<br>Leu        | CTA<br>Leu        | TGT<br>Cys<br>430 | GTA<br>Val        | AGA<br>Arg        | • | 1296 |
| GGG<br>Gly        | ATA<br>Ile        | ATA<br>Ile<br>435 | ACT<br>Thr        | TCT<br>Ser        | AAA<br>Lys        | ACT<br>Thr        | AAA<br>Lys<br>440 | TCA<br>Ser        | TTA<br>Leu        | GAT<br>Asp        | AAA<br>Lys        | GGA<br>Gly<br>445 | TAC<br>Tyr        | AAT<br>Asn        | AAG<br>Lys        |   | 1344 |
| Ala               | Leu<br>450        | Asn               | Asp               | Leu               | Cys               | 11e<br>455        | Lys               | Val               | Asn               | Asn               | Trp<br>460        | Asp               | Leu               | TTT<br>Phe        | Phe               | • | 1392 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | GAA<br>Glu        |                   |   | 1440 |
| Íle               | Thr               | Ser               | Asp               | Thr<br>485        | Asn               | Ile               | Glu               | Ala               | Ala<br>490        | Glu               | Glu               | Asn               | Ile               | AGT<br>Ser<br>495 | Leu               |   | 1488 |
| Asp               | Leu.              | Ile               | Gln<br>500        | Gln               | Tyr               | Tyr               | Leu               | Thr<br>505        | Phe               | Asn               | Phe               | Asp               | Asn<br>510        | GAA<br>Glu        | Pro               | 7 | 1536 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | CAA<br>Gln        |                   |   | 1584 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | _                 |                   | TAT               |                   |   | 1632 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | TTT<br>Phe        |                   |   | 1680 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | GCA<br>Ala<br>575 |                   |   | 1728 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | GTA<br>Val        |                   |   | 1776 |
|                   |                   |                   | Lys               |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | GTA<br>Val        |                   |   | 1824 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | ACT<br>Thr        |                   |   | 1872 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | CCT<br>Pro        |                   |   | 1920 |

| TT<br>Le          | A AA<br>u As      | T AI<br>n Il      | A GO<br>e Gl       | T AA<br>y As<br>64    | n me              | G TT/<br>t Lev        | ч ТАЗ               | r AA              | A GA:<br>s As <sub>1</sub><br>650 | aA q              | T TT                | r gra<br>e Val        | A GGT                 | GC<br>Al<br>65    | T TTA<br>a Leu<br>5   | . 1968 |
|-------------------|-------------------|-------------------|--------------------|-----------------------|-------------------|-----------------------|---------------------|-------------------|-----------------------------------|-------------------|---------------------|-----------------------|-----------------------|-------------------|-----------------------|--------|
| AT.               | A TT<br>e Ph      | T TC<br>e Se      | A GG<br>r Gl<br>66 | A WI                  | T GT<br>a Va      | T ATT                 | CTC<br>Leu          | TTI<br>Let<br>669 | ı Glu                             | A TT              | T ATA               | A CCA                 | A GAC<br>O Glu<br>670 | ı Il              | T GCA<br>e Ala        | 2016   |
| AT                | A CC              | T GT.             | T re               | A GG<br>u Gl          | T ACT             | r TTT                 | GCA<br>Ala<br>680   | Let               | r GTA<br>ı Val                    | A TC              | A TAT               | ATT                   | • Ala                 | AA S              | T AAG<br>n Lys        | 2064   |
| GT:<br>Val        | CT)<br>Let<br>690 | 1 Th              | C GT<br>r Va       | T CAL                 | A ACA             | A ATA<br>Ile<br>695   | GAT<br>Asp          | AAT<br>Asn        | GCT<br>Ala                        | TTA<br>Lev        | AGT<br>Ser<br>700   | Lys                   | AGA<br>Arg            | AA:<br>Asi        | r GAA<br>n Glu        | 2112   |
| Lys<br>705        | til               | GA:               | r GAG              | G GTO                 | TAT<br>Tyr<br>710 | rys                   | TAT<br>Tyr          | ATA<br>Ile        | GTA<br>Val                        | ACA<br>Thr<br>715 | Asn                 | TGG<br>Trp            | TTA<br>Leu            | GCA<br>Ala        | A AAG<br>A Lys<br>720 | 2160   |
| GT1<br>Val        | AAT<br>Asr        | ACA<br>Thi        | CAC<br>CGl         | 3 ATT<br>1 Ile<br>725 | : Asp             | CTA<br>Leu            | ATA<br>Ile          | AGA<br>Arg        | AAA<br>Lys<br>730                 | Lys               | ATG<br>Met          | AAA<br>Lys            | GAA<br>Glu            | GCT<br>Ala<br>735 | TTA<br>Leu            | 2208   |
| GAA<br>Glu        | AAT<br>Asn        | CAA<br>Glr        | 740                | A GAA<br>a Glu        | GCA<br>Ala        | ACA<br>Thr            | AAG<br>Lys          | GCT<br>Ala<br>745 | ATA<br>Ile                        | ATA<br>Ile        | AAC<br>Asn          | TAT<br>Tyr            | CAG<br>Gln<br>750     | TAT               | AAT<br>Asn            | 2256   |
| CAA<br>Gln        | TAT               | ACT<br>Thr<br>755 | . GIU              | GAA<br>Glu            | GAG<br>Glu        | AAA<br>Lys            | AAT<br>Asn<br>760   | AAT<br>Asn        | ATT<br>Ile                        | AAT<br>Asn        | TTT<br>Phe          | AAT<br>Asn<br>765     | ATT<br>Ile            | GAT<br>Asp        | GAT<br>Asp            | 2304   |
| TTA<br>Leu        | AGT<br>Ser<br>770 | Ser               | Lys                | CTT<br>Leu            | AAT<br>Asn        | GAG<br>Glu<br>775     | TCT<br>Ser          | ATA<br>Ile        | AAT<br>Asn                        | AAA<br>Lys        | GCT<br>Ala<br>780   | ATG<br>Met            | ATT<br>Ile            | AAT<br>Asn        | ATA<br>Ile            | 2352   |
| AAT<br>Asn<br>785 | AAA<br>Lys        | TTT               | TTG<br>Leu         | AAT<br>Asn            | CAA<br>Gln<br>790 | TGC<br>Cys            | TCT<br>Ser          | GTT<br>Vaļ        | TCA<br>Ser                        | TAT<br>Tyr<br>795 | TTA<br>Leu          | ATG<br>Met            | AAT<br>Asn            | TCT<br>Ser        | ATG<br>Met<br>800     | 2400   |
| ATC<br>Ile        | CCT<br>Pro        | TAT<br>Tyr        | GGT<br>Gly         | GTT<br>Val<br>805     | AAA<br>Lys        | CGG<br>Arg            | TTA<br>Leu          | GAA<br>Glu        | GAT<br>Asp<br>810                 | TTT<br>Phe        | GAT<br>Asp          | GCT<br>Ala            | AGT<br>Ser            | CTT<br>Leu<br>815 | AAA<br>Lys            | 2448   |
| GAT<br>Asp        | GCA<br>Ala        | TTA<br>Leu        | TTA<br>Leu<br>820  | AAG<br>Lys            | TAT<br>Tyr        | ATA<br>Ile            | Tyr                 | GAT<br>Asp<br>825 | AAT<br>Asn                        | AGA<br>Arg        | GGA<br>Gly          | Thr                   | TTA<br>Leu<br>830     | ATT<br>Ile        | GGT<br>Gly            | 2496   |
| CAA<br>Gln        | GTA<br>Val        | GAT<br>Asp<br>835 | AGA<br>Arg         | TTA<br>Leu            | AAA<br>Lys        | Asp :                 | AAA (<br>Lys<br>840 | GTT<br>Val        | AAT .<br>Asn .                    | TAA<br>Nan        | Thr                 | CTT .<br>Leu .<br>845 | AGT :<br>Ser :        | ACA<br>Thr        | GAT<br>Asp            | 2544   |
| ATA<br>Ile        | CCT<br>Pro<br>850 | TTT<br>Phe        | CAG<br>Gln         | CTT<br>Leu            | Ser               | AAA '<br>Lys '<br>855 | IAC (<br>Iyr \      | GTA (<br>Val ,    | GAT Asp                           | Asn               | CAA<br>Gln i<br>860 | AGA :<br>Arg :        | TTA :<br>Leu l        | ITA<br>Leu        | TCT<br>Ser            | 2592   |
|                   |                   |                   |                    | TAT<br>Tyr            |                   |                       | ,                   |                   |                                   |                   |                     |                       |                       |                   |                       | 2613   |

# (2) INFORMATION FOR SEQ ID NO: 8:

<sup>(</sup>i) SEQUENCE CHARACTERISTICS:

<sup>(</sup>A) LENGTH: 871 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Pro Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr 135 Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr 170 Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Asn 280 Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val 295 Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys

Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu

Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp 340

Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn 355

Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr 370 380

Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn 385 390 395 400

Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu 405 410 415

Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg

Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys
435
440
445

Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe 450 455 460

Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu 465 470 475 480

Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu
485 490 495

Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro 500 505 510

Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu 515 525

Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu 530 540

Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu 545 550 550

His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu 565 570 575

Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys 580 585

Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu
595 600 605

Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr 610 620

Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala 625 630 635 640

Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu 645 650 655

Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala 660 665 670

Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys 675 680 685.

| Val                           | Leu<br>690                 | Thr   | Val  | Gln   | Thr   | Ile<br>695   | Asp  | Asn   | Ala                            | Leu                             | Ser<br>700                      | Lys               | Arg                | Asn                            | Glu               |    | •          |
|-------------------------------|----------------------------|---|--|---|---|--|--|---|--------------------------------|---------------------------------|---------------------------------|-------------------|--------------------|--------------------------------|-------------------|----|------------|
| Lys<br>705                    | Trp                        | Asp   | Glu  | Val   | Tyr<br>710  | Lys  | Tyr  | Ile   | Val                            | Thr<br>715                      |                                 | Trp               | Leu                | Ala                            | Lys<br>720        |    |            |
| Val                           | Asn                        | Thr   | Gln  | 11e<br>725  | Asp   | Leu  | Ile  | Arg   | Lys<br>730                     | Lys                             | Met                             | Lys               | Glu                | Ala<br>735                     | Leu               |    |            |
| Glu                           | Asn                        | Gln   | Ala<br>740   | Glu   | Ala   | Thr  | Lys  | Ala<br>745  | Ile                            | Ile                             | Asn                             | Tyr               | Gln<br>750         | Tyr                            | Asn               |    |            |
| Gln                           | Tyr                        | Thr<br>755  | Glu  | Glu   | Glu   | Lys  | Asn<br>760   | Asn   | Ile                            | Asn                             | Phe                             | Asn<br>765        | Ile                | Asp                            | Asp               |    |            |
| Leu                           | Ser<br>770                 | Ser   | Lys  | Leu   | Asn   | Glu<br>775   | Ser  | Ile   | Asn                            | Lys                             | Ala<br>780                      | -                 | Ile                | Asn                            | Ile               | •  |            |
| Asn<br>785                    | Lys                        | Phe   | Leu  | Asn   | Gln<br>790  | Cys  | Ser  | Val   | Ser                            | Tyr<br>795                      | Leu                             | Met               | Asn                | Ser,                           | Met<br>800        |    |            |
|                               | Pro                        |   |  | 805   |   |  |  |   | 810                            |                                 | -                               |                   |                    | 815                            | •                 |    |            |
|                               | Ala                        | • :   | 820  |   |   |  |  | 825   |                                |                                 |                                 | ٠.                | 830                |                                |                   |    |            |
|                               | Val                        | 835   |  |   | -   | •  | 840  |   |                                |                                 | •                               | 845               | •                  |                                | _                 |    |            |
|                               | Pro<br>850                 |   | •  |   |   | 855  | Tyr  | Val   | Asp                            | Asn                             | Gln<br>860                      | Arg               | Leu                | Leu                            | Ser               | ٠. |            |
|                               |                            |   |  |   |   |  |  |   |                                |                                 |                                 |                   |                    |                                |                   |    |            |
| 865                           |                            |   | Glu  | _   | 870   |  | ·.   | ,   |                                |                                 |                                 |                   | •                  |                                | •                 |    |            |
|                               | INFO                       | ORMA:   | rion   | FOR   | 870<br>SEQ  | ID N   |  | •   |                                | . ·<br>:                        | ٠                               |                   |                    | ,                              | •                 | ,  |            |
| 865                           | INFO                       | ORMA:<br>SE(  | CION<br>QUENC  | FOR<br>CE CH  | 870<br>SEQ<br>IARAC<br>I: 26  | ID N<br>TERI   | STIC   | CS:<br>pair   | s                              | . ·                             |                                 | · .               |                    |                                |                   | ,  |            |
| 865                           | INFO                       | ORMA:<br>SE(<br>()<br>(1  | OUENC<br>CION  | FOR<br>CE CH<br>ENGTH<br>(PE:<br>CRANI  | 870<br>SEQ<br>IARAC<br>I: 26<br>nucl                                      | ID N<br>TERI<br>28 b<br>eic<br>ESS:                        | STIC<br>ase<br>acid<br>doub  | CS:<br>pair   | s                              | :                               |                                 |                   |                    |                                | •                 |    |            |
| 865                           | INFO                       | ORMA:<br>SE(<br>()<br>()<br>()  | TION  OUENCE  A) LE  B) TY  C) ST  | FOR<br>CE CH<br>ENGTH<br>(PE:<br>TRANII<br>OPOLO                                    | 870<br>SEQ<br>IARAC<br>I: 26<br>nucl<br>EDNE<br>GY:                       | ID N<br>TERI<br>28 k<br>eic<br>SS:<br>line                 | STIC<br>ase<br>acid<br>doub<br>ar                                      | CS:<br>paix<br>d                                    |                                |                                 |                                 |                   | ,                  |                                |                   |    |            |
| 865                           | [NFC (i)                   | SEC<br>(1)<br>(1)<br>(1)<br>(1)<br>(1)<br>(2)<br>(1)                              | TION  QUENC  A) LE  B) TY  C) ST  C) TC  | FOR CE CHENGTH (PE: CRANIC CPOLO LE TY  LE TY  AME/H                                | 870 SEQ HARACI: 26 nucl DEDNE OGY: (PE:                                   | ID N<br>TERI<br>28 t<br>eic<br>SS:<br>line<br>DNA          | STIC<br>pase<br>acid<br>doub<br>ar<br>(ger                             | CS:<br>paix<br>d                                    |                                |                                 |                                 |                   |                    |                                |                   |    |            |
| 865                           | (ii)<br>(ii)<br>(ix)       | SEX<br>(I<br>(I<br>(I<br>MOI<br>FEZ<br>(I   | CION  QUENCA) LE  B) TY  C) ST  C) TC  LECUI  ATURE  | FOR CE CHENGTH (PE: CRANIC CPOLO LE TY AME/H CCATI                                  | 870 SEQ HARACI: 26 nucl DEDNE OGY: (PE:                                   | ID N TTERI 28 h eic ESS: line DNA CDS                      | STIC<br>Dase<br>acid<br>doub<br>ar<br>(ger                             | CS:<br>pair<br>pair<br>pole                         | <b>:)</b>                      | ): <b>9</b> :                   |                                 |                   |                    |                                |                   |    |            |
| 865<br>(2)                    | (ii)<br>(ii)<br>(ix)       | SE() (I) (I) (I) (I) (I) (I) (I) (I) (I) (  | CUENCE CTG   | FOR CE CHENGTH (PE: CRANIC CPOLO LE TY CAME/H CCATI                                 | 870 SEQ HARACI : 26 nucl DEDNE DGY: (PE: CEY: CON:1                       | ID N TERI 28 h eic ESS: line DNA CDS26                     | STIC   | CS: pair pair ple comic                             | :)<br>ID NO                    | AAG                             | GAC                             |                   |                    |                                |                   |    | 4.8        |
| ATG<br>Met<br>1               | (ii) (ix) (xi) CAG         | SE(<br>()<br>()<br>()<br>()<br>()<br>()<br>()<br>()<br>()<br>()<br>()<br>()<br>() | CION  QUENC  ATURE  ATURE  ATURE  ATURE  OUENC  GTG  Val   | FOR CE CHENGTH (PE: CRANIC) CO CAT CE CHENGTH CO CAT CE CHENGTH AAC AAC AAC AAC TAC | 870 SEQ HARACI 1: 26 nucl DEDNE OGY: CPE: CON: 1 ESCRI AAG Lys            | ID N CTERI C28 b eic CSS: line DNA CDS CPTIC CAG Gln       | STIC<br>Phe  | Pair<br>pair<br>ole<br>nomic<br>SEQ 1<br>AAC<br>Asn | TAT<br>Tyr<br>10               | AAG<br>Lys<br>GCC               | GAC<br>Asp<br>GGC               | Pro<br>CAG        | Val<br>ATG         | Asn<br>15<br>CAG               | Gly               |    | <b>4</b> 8 |
| ATG<br>Met<br>1<br>GTT<br>Val | (ii) (ix) (xi) CAG Gln GAC | SE( (I) (I) (I) (I) (I) (I) (I) (I) (I) (I  | CUENCE COLOR | FOR CE CHENGTH (PE: CRANIC OPOLO LE TY AME/H OCATI CE DE AAC ASn 5 TAC Tyr          | 870 SEQ HARACH: 26 nucl DEDNE DGY: (PE: CEY: CON: 1 ESCRI AAG Lys ATC Ile | ID N TERI 28 t eic SS: line DNA CDS26 PTIC CAG Gln AAA Lys | STIC<br>doub<br>ar<br>(ger<br>528<br>ON: S<br>TTC<br>Phe<br>ATT<br>Ile | Pro 25  | TAT<br>Tyr<br>10<br>AAC<br>ABD | AAG<br>Lys<br>GCC<br>Ala<br>TGG | GAC<br>Asp<br>GGC<br>Gly<br>GTT | Pro<br>CAG<br>Gln | Val ATG Met 30 CCG | Asn<br>15<br>CAG<br>Gln<br>GAA | Gly<br>CCG<br>Pro |    |            |

| GC<br>Al<br>6          | a na       | G CA<br>s Gl | NG G1<br>.n Va     | G CC               | A GT:<br>o Val        | r Sei          | A TAC        | C TA       | C GA<br>r As          | T TC<br>p Se<br>7 | r Th              | C TA         | T CI               | G AC              | GC ACA<br>r Thr<br>80 |                  |
|------------------------|------------|--------------|--------------------|--------------------|-----------------------|----------------|--------------|------------|-----------------------|-------------------|-------------------|--------------|--------------------|-------------------|-----------------------|------------------|
| GA(<br>As <sub>1</sub> | C AA       | C GA<br>n Gl | G AA<br>u Ly       | G GA<br>'s As<br>8 | D WEI                 | TAC<br>Tyi     | CTC<br>Lev   | AAI<br>Ly: | G GG<br>S Gl          | y Va              | G AC<br>l Th      | C AA<br>r Ly | A TT<br>s Le       | u Ph              | C GAG<br>e Glu<br>5   | 288              |
| CG                     | TAT        | TA<br>Ty     | T TC<br>r Se<br>10 | rin                | T GAC                 | CTC<br>Leu     | GGC<br>Gly   | CG:<br>Arg | g Met                 | G CT              | G CT              | G AC         | C TC<br>r Se<br>11 | r Il              | C GTC<br>e Val        | 336              |
| Arg                    | 3 GTZ      | 11           | e Pr               | o Pne              | e Trp                 | GIY            | 120          | ser        | . Thi                 | r Ile             | e Ası             | 125          | r Gli<br>5         | u Le              | G AAG<br>u Lys        | 384              |
| GT1<br>Val             | 130        | : AS         | C AC<br>p Th       | T AA(<br>r Asi     | TGC<br>Cys            | Ile<br>135     | Asn          | GTC<br>Val | ATC<br>Ile            | CAA<br>Glr        | CCA<br>Pro<br>140 | ) Ası        | GG;<br>Gly         | r ag<br>/ Se:     | TAC<br>Tyr            | 432              |
| 145                    | Sei        | GI           | . GI               | ı Lev              | 150                   | ren            | Val          | Ile        | · Ile                 | 155               | Pro               | Ser          | : Ala              | Asi               | ATT<br>Ile<br>160     | 480              |
| TTE                    | GIN        | Pne          | e GIt              | 165                | rys                   | Ser            | Phe          | Gly        | 170                   | Glu               | Val               | Leu          | Asn                | 175               |                       | 528              |
| Arg                    | ASN        | GT           | 180                | GIY                | Ser                   | Tnr            | GIn          | Tyr<br>185 | Ile                   | Arg               | Phe               | Ser          | Pro<br>190         | Asp               | Phe                   | <b>576</b>       |
| Inr                    | Pne        | 195          | Pne                | GIU                | GAG<br>Glu            | ser            | 200          | GIn        | Val                   | qaA               | Thr               | Asn<br>205   | Pro                | Leu               | Leu                   | 624              |
| GIĀ                    | 210        | GIÀ          | гÀз                | Pne                | GCA<br>Ala            | Thr<br>215     | Asp          | Pro        | Ala                   | Val               | Thr<br>220        | Leu          | Ala                | His               | Glu                   | 672              |
| 225                    | ire        | nls          | Ala                | GIĀ                | CAT<br>His<br>230     | Arg            | Leu          | Tyr        | Gly                   | 11e<br>235        | Ala               | Ile          | Asn                | Pro               | Asn<br>240            | . 720            |
| Arg                    | VaI        | Pne          | Lys                | Val<br>245         | AAC<br>Asn            | Thr            | Asn .        | Ala        | Tyr<br>250            | Tyr               | Glu               | Met          | Ser                | Gly<br>255        | Leu                   | 768              |
| GIU                    | Val        | Ser          | 260                | Glu                | GAA (<br>Glu          | Leu .          | Arg (        | Thr<br>265 | Phe                   | Gly               | Gly               | His          | Asp<br>270         | Ala               | Lys                   | 816              |
| Pne                    | ITE        | 275          | ser                | Leu                | CAG (<br>Gln (        | Glu :          | Asn (<br>280 | Glu        | Phe                   | Arg               | Leu               | Tyr<br>285   | Tyr                | Tyr               | Asn                   | 864              |
| ràs                    | Phe<br>290 | Lys          | Asp                | Ile                | •                     | Ser :<br>295   | Thr 1        | Leu .      | Asn                   | Lys               | Ala<br>300        | Lys          | Ser                | Ile               | Val                   | 912              |
| 305                    | Thr        | Thr          | Ala                | Ser                | TTA (<br>Leu (<br>310 | 3ln 7          | Cyr N        | let :      | Lys .                 | Asn<br>315        | Val               | Phe          | Lys                | Glu               | Lys<br>320            | 960 <sub>.</sub> |
| TAT (<br>Tyr )         | CTC<br>Leu | CTA<br>Leu   | Ser                | GAA<br>Glu<br>325  | GAT A                 | ACA 7<br>Thr S | CT C         | Sly 1      | AAA '<br>Lys '<br>330 | TTT 'Phe          | TCG<br>Ser        | GTA Val      | Asp :              | AAA<br>Lys<br>335 | TTA<br>Leu            | 1008             |

|   |                            |                   |                   |                   |            |            |                   |                   |            |            |                   |            |                     |            | 1           |                    |     |      |
|---|----------------------------|-------------------|-------------------|-------------------|------------|------------|-------------------|-------------------|------------|------------|-------------------|------------|---------------------|------------|-------------|--------------------|-----|------|
|   | AAA<br>Lys                 | TTT<br>Phe        | GAT<br>Asp        | AAC<br>Lys<br>340 | , rer      | TAC<br>Tyr | Lys               | A ATO             | 345        | Thi        | A GAC             | AT         | TAC<br>Tyr          | ACA<br>Thr | Glu         | GAT<br>Asp         |     | 1056 |
|   | AAT<br>Asn                 | TTI<br>Phe        | GT1<br>Val<br>355 | . Lys             | TTT<br>Phe | TTI<br>Phe | AAA<br>Lys        | GTA<br>Val<br>360 | . Lev      | AAC<br>Asr | AGA<br>Arg        | AA/<br>Lys | A ACA<br>Thr<br>365 | Tyr        | TTC         | AAT<br>Asn         | *   | 1104 |
|   | TTT                        | GAT<br>Asp<br>370 | . råa             | GCC               | GTA<br>Val | Phe        | AAG<br>Lys<br>375 | Ile               | AAT<br>Asn | ATA        | GTA<br>Val        | CCI<br>Pro | Lys                 | GTA<br>Val | AAT<br>Asn  | TAC                |     | 1152 |
|   | ACA<br>Thr<br>385          | TIE               | TAT<br>Tyr        | GAT<br>Asp        | GGA<br>Gly | Phe        | Asn               | TTA<br>Leu        | AGA<br>Arg | AAT<br>Asn | ACA<br>Thr<br>395 | AAT<br>Asn | TTA<br>Leu          | GCA<br>Ala | GCA<br>Ala  | AAC<br>Asn<br>400  |     | 1200 |
|   | Pne                        | Asn               | GIY               | GIn               | 405        | Thr        | Glu               | Ile               | Asn        | Asn<br>410 | Met               | Asn        | Phe                 | Thr        | Lys<br>415  | CTA<br>Leu         | ٠   | 1248 |
|   | Lys                        | Asn               | Phe               | 420               | GIA        | Leu        | Phe               | Glu               | Phe<br>425 | Tyr        | Lys               | Leu        | Leu                 | Cys<br>430 | <u>V</u> al |                    |     | 1296 |
|   | GGG<br>Gly                 | ATA               | ATA<br>Ile<br>435 | ACT<br>Thr        | TCT<br>Ser | AAA<br>Lys | ACT<br>Thr        | AAA<br>Lys<br>440 | TCA<br>Ser | TTA<br>Leu | GAT<br>Asp        | AAA<br>Lys | GGA<br>Gly<br>445   | TAC<br>Tyr | AAT<br>Asn  | AAG<br>Lys         |     | 1344 |
|   | Ser                        | 450               | Asp               | GIA               | Ala        | Leu        | 455               | Asp               | Leu        | Cys        | Ile               | Lys<br>460 | GTT<br>Val          | Asn        | Asn         | Trp                | -   | 1392 |
|   | <b>A</b> 8p<br><b>4</b> 65 | Leu               | Phe               | Phe               | .ser       | 470        | Ser               | Glu               | Asp        | Asn        | Phe 475           | Thr        | AAT<br>Asn          | Asp        | Leu         | Asn<br>480         |     | 1440 |
|   | гув                        | GIA               | Glu               | Glu               | 11e<br>485 | Thr        | Ser               | Asp               | Thr        | Asn<br>490 | Ile               | Glu        | GCA<br>Ala          | Ala        | Glu<br>495  | Glu                | *   | 1488 |
|   | Asn                        | 11e               | ser               | <b>Leu</b> 500    | Asp        | Leu        | Ile               | Gln               | Gln<br>505 | Tyr        | Tyr               | Leu        | Thr                 | Phe<br>510 | Asn         |                    |     | 1536 |
|   | Asp                        | Asn               | 515               | Pro               | Glu        | Asn        | Ile               | Ser<br>520        | Ile        | Glu        | Asn               | Leu        | TCA<br>Ser<br>525   | Ser        | Asp         | Ile                |     | 1584 |
|   | ile                        | 530               | Gin               | Leu               | Glu        | Leu        | Met<br>535        | Pro               | Asn        | Ile        | Glu               | Arg<br>540 | TTT<br>Phe          | Pro        | Asn         | Gly                | -)- | 1632 |
| • | Lys<br>545                 | rys               | Tyr               | Glu               | Leu        | Asp<br>550 | Lys               | Tyr               | Thr        | Met        | Phe<br>555        | His        | TAT<br>Tyr          | Leu        | Arg         | <b>A</b> la<br>560 |     | 1680 |
| • | GIN                        | Glu               | Pne               | Glu               | H18<br>565 | Gly        | Lys               | Ser               | Arg        | 11e<br>570 | Ala               | Leu        | ACA .<br>Thr        | Asn        | Ser<br>575  | Val                |     | 1728 |
| 4 | Asn                        | Glu               | Ala               | <b>Leu</b> 580    | Leu        | Asn        | Pro               | Ser               | Arg<br>585 | Val        | Tyr               | Thr        |                     | Phe<br>590 | Ser         | Ser                |     | 1776 |
|   | GAC<br>Asp                 | TAT<br>Tyr        | GTA<br>Val<br>595 | AAG<br>Lys        | AAA<br>Lys | GTT<br>Val | Asn               | AAA<br>Lys<br>600 | GCT<br>Ala | ACG<br>Thr | GAG<br>Glu        | GCA<br>Ala | GCT Ala 1           | ATG<br>Met | TTT<br>Phe  | TTA<br>Leu         |     | 1824 |
|   |                            |                   |                   |                   |            |            |                   |                   |            |            |                   |            |                     |            |             |                    |     |      |

| GGC<br>Gly        | TGG<br>Trp<br>610  | GTA<br>Val        | GAA<br>Glu        | CAA<br>Gln        | TTA<br>Leu        | GTA<br>Val<br>615 | TAT<br>Tyr        | GAT<br>qaa        | TTT<br>Phe        | ACC<br>Thr        | GAT<br>Asp<br>620 | GAA<br>Glu        | ACT<br>Thr        | AGC<br>Ser        | GAA<br>Glu        | • | 1872 |
|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---|------|
| GTA<br>Val<br>625 | AGT                | ACT<br>Thr        | ACG<br>Thr        | GAT<br>Asp        | AAA<br>Lys<br>630 | ATT<br>Ile        | GCG<br>Ala        | GAT<br>Asp        | ATA<br>Ile        | ACT<br>Thr<br>635 | ATA<br>Ile        | ATT<br>Ile        | ATT<br>Ile        | CCA<br>Pro        | TAT<br>Tyr<br>640 |   | 1920 |
| ΔΤΔ               | GGA<br>Gly         | CCT<br>Pro        | GCT<br>Ala        | TTA<br>Leu<br>645 | AAT<br>Asn        | ATA<br>Ile        | GGT<br>Gly        | AAT<br>Asn        | ATG<br>Met<br>650 | TTA<br>Leu        | TAT<br>Tyr        | AAA<br>Lys        | GAT<br>Asp        | GAT<br>Asp<br>655 | TTT<br>Phe        |   | 1968 |
| GTA<br>Val        | GGT<br>Gly         | GCT<br>Ala        | TTA<br>Leu<br>660 | Ile               | TTT               | TCA<br>Ser        | GGA<br>Gly        | GCT<br>Ala<br>665 | GTT<br>Val        | ATT<br>Ile        | CTG<br>Leu        | TTA<br>Leu        | GAA<br>Glu<br>670 | TTT               | ATA<br>Ile        |   | 2016 |
| CCA<br>Pro        | <b>GA</b> G<br>Glu | ATT<br>Ile<br>675 | GCA<br>Ala        | ATA<br>Ile        | CCT<br>Pro        | GTA<br>Val        | TTA<br>Leu<br>680 | GGT<br>Gly        | ACT<br>Thr        | TTT<br>Phe        | GCA<br>Ala        | CTT<br>Leu<br>685 | Val               | TCA<br>Ser        | TAT               | - | 2064 |
| ATT<br>Ile        | GCG<br>Ala<br>690  | AAT<br>Asn        | AAG<br>Lys        | GTT<br>Val        | CTA<br>Leu        | ACC<br>Thr<br>695 | GTT<br>Val        | CAA<br>Gln        | ACA<br>Thr        | ATA<br>Ile        | GAT<br>Asp<br>700 | AAT<br>Asn        | GCT<br>Ala        | TTA<br>Leu        | AGT<br>Ser        |   | 2112 |
| AAA<br>Lys<br>705 | AGA<br>Arg         | TAA<br>Asn        | GAA<br>Glu        | AAA<br>Lys        | TGG<br>Trp<br>710 | GAT<br>Asp        | GAG<br>Glu        | GTC<br>Val        | TAT<br>Tyr        | AAA<br>Lys<br>715 | TAT<br>Tyr        | ATA<br>Ile        | GTA<br>Val        | ACA<br>Thr        | AAT<br>Asn<br>720 |   | 2160 |
| TGG<br>Trp        | TTA<br>Leu         | GCA<br>Ala        | Lys               | GTT<br>Val<br>725 | AAT<br>Asn        | ACA<br>Thr        | CAG<br>Gln        | Ile               | GAT<br>Asp<br>730 | CTA<br>Leu        | ATA<br>Ile        | AGA<br>Arg        | AAA<br>Lys        | AAA<br>Lys<br>735 | ATG<br>Met        |   | 2208 |
| AAA<br>Lys        | GAA<br>Glu         | GCT<br>Ala        | TTA<br>Leu<br>740 | GAA<br>Glu        | AAT<br>Asn        | CAA<br>Gln        | GCA<br>Ala        | GAA<br>Glu<br>745 | GCA<br>Ala        | ACA<br>Thr        | AAG<br>Lys        | GCT<br>Ala        | ATA<br>Ile<br>750 | ATA<br>Ile        | AAC<br>Asn        |   | 2256 |
| TAT               | CAG<br>Gln         | TAT<br>Tyr<br>755 | AAT<br>Asn        | CAA<br>Gln        | TAT<br>Tyr        | ACT<br>Thr        | GAG<br>Glu<br>760 | GAA<br>Glu        | GAG<br>Glu        | AAA<br>Lys        | AAT<br>Asn        | AAT<br>Asn<br>765 | ATT               | AAT<br>Asn        | TTT<br>Phe        |   | 2304 |
| AAT<br>Asn        | ATT<br>Ile<br>770  | GAT<br>Asp        | GAT<br>Asp        | TTA<br>Leu        | AGT<br>Ser        | TCG<br>Ser<br>775 | AAA<br>Lys        | CTT<br>Leu        | AAT<br>Asn        | GAG<br>Glu        | TCT<br>Ser<br>780 | ATA<br>Ile        | AAT<br>Asn        | AAA<br>Lys        | GCT<br>Ala        |   | 2352 |
| ATG<br>Met<br>785 | ATT<br>Ile         | AAT<br>Asn        | ATA<br>Ile        | AAT<br>Asn        | AAA<br>Lys<br>790 | TTT               | TTG<br>Leu        | AAT<br>Asn        | CAA<br>Gln        | TGC<br>Cys<br>795 | TCT<br>Ser        | GTT<br>Val        | TCA<br>Ser        | TAT<br>Tyr        | TTA<br>Leu<br>800 |   | 2400 |
| ATG<br>Met        | AAT<br>Asn         | TCT<br>Ser        | ATG<br>Met        | ATC<br>Ile<br>805 | Pro               | TAT<br>Tyr        | GGT<br>Gly        | GTT<br>Val        | AAA<br>Lys<br>810 | CGG               | TTA<br>Leu        | GAA<br>Glu        | gat<br>Asp        | TTT<br>Phe<br>815 | GAT<br>Asp        |   | 2448 |
| GCT<br>Ala        | AGT<br>Ser         | CTT<br>Leu        | AAA<br>Lys<br>820 | gat<br>Asp        | GCA<br>Ala        | TTA<br>Leu        | TTA<br>Leu        | AAG<br>Lys<br>825 | TAT               | ATA<br>Ile        | TAT<br>Tyr        | GAT<br>Asp        | AAT<br>Asn<br>830 | AGA<br>Arg        | GGA<br>Gly        |   | 2496 |
| ACT<br>Thr        | TTA<br>Leu         | ATT<br>Ile<br>835 | GGT<br>Gly        | CAA<br>Gln        | GTA<br>Val        | GAT<br>Asp        | AGA<br>Arg<br>840 | TTA<br>Leu        | AAA<br>Lys        | ĠAT<br>Asp        | AAA<br>Lys        | GTT<br>Val<br>845 | AAT<br>Asn        | TAA<br>neA        | ACA<br>Thr        |   | 2544 |
| CTT               | AGT<br>Ser<br>850  | ACA<br>Thr        | GAT<br>Asp        | ATA<br>Ile        | CCT<br>Pro        | TTT<br>Phe<br>855 | CAG<br>Gln        | CTT               | TCC<br>Ser        | TAS               | TAC<br>Tyr<br>860 | GTA<br>Val        | GAT<br>Asp        | TAA<br>Asn        | CAA<br>Gln        |   | 2592 |
| AGA<br>Arg<br>865 | TTA<br>Leu         | TTA<br>Leu        | TCT<br>Ser        | ACA<br>Thr        | TTT<br>Phe<br>870 | ACT<br>Thr        | GAA<br>Glu        | TAT               | ATT               | AAG<br>Lys<br>875 | TAA<br>*          |                   |                   |                   |                   |   | 2628 |

- (2) INFORMATION FOR SEQ ID NO: 10:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 876 amino acids (B) TYPE: amino acid

    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly 10

Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro

Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg

Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu

Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr

Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu

Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val

Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys 115 120 125

Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr

Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile 145 150 155 160

Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr

Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe

Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu

Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu

Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn

Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu

Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys

Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Asn 280

Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val 295

Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Ph Ser Val Asp Lys Leu 330 Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn 395 Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg 425 Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys 440 Ser Ala Asp Gly Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu 490 Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe 505 Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe-Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe

48

96



Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr . 675 Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn 705 710 715 720 Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp 805 Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr 835 Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln 860

- (2) INFORMATION FOR SEQ ID NO: 11:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 2637 base pairs
    - (B) TYPE: nucleic acid

Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys 870

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE\_TYPE: DNA (genomic)
- (ix) FEATURE:

  - (A) NAME/KEY: CDS
    (B) LOCATION:1..2637
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATG CAG TTC GTG AAC AAG CAG TTC AAC TAT AAG GAC CCT GTA AAC GGT Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly

GTT GAC ATT GCC TAC ATC AAA ATT CCA AAC GCC GGC CAG ATG CAG CCG Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro 20 25



| 35  | : bys lie mis                         | 40                       | TC TGG GTT ATT CO<br>le Trp Val Ile Pr<br>45  | o Glu Arg              |
|---|---------------------------------------|--------------------------|---|------------------------|
| 50  | 55                                    | GIU GIY AS               | AC TTG AAC CCG CC<br>SP Leu Asn Pro Pr<br>60  | o Pro Glu              |
| 65  | 70                                    | Tyr Tyr As               | T TCA ACC TAT CT<br>p Ser Thr Tyr Le<br>75    | u Ser Thr<br>80        |
| ASP ASH GIU LYS                               | 85                                    | Dea Lys GI               | A GTG ACC AAA TT<br>y Val Thr Lys Le<br>0     | u Phe Glu<br>95        |
| arg fie fyr ser<br>100                        | ini Asp Leu                           | 105                      | G CTG CTG ACC TC<br>t Leu Leu Thr Se<br>11    | r Ile Val              |
| 115   | Phe Trp Gly                           | 120                      | C ATT GAC ACG GAG<br>r Ile Asp Thr Glu<br>125 | Leu Lys                |
| 130   | 135                                   | Asn val Ile              | 140   | Ser Tyr                |
| AGA TCT GAA GAA<br>Arg Ser Glu Glu<br>145     | 150                                   | val lie lle              | e Gly Pro Ser Ala<br>155                      | Asp Ile<br>160         |
|   | cys Lys Ser 1<br>165                  | Phe Gly His<br>170       | Glu Val Leu Asn                               | Leu Thr<br>175         |
| CGT AAC GGT TAC Arg Asn Gly Tyr               | Gly Ser Inr G                         | 185                      | Arg Phe Ser Pro<br>190                        | Asp Phe                |
| ACG TTC GGT TTC ( Thr Phe Gly Phe ( 195       | or Gru Ser L                          | eu Glu Val               | Asp Thr Asn Pro<br>205                        | Leu Leu                |
| GGT GCA GGC AAG T<br>Gly Ala Gly Lys I<br>210 | ne Ala Inr A<br>215                   | sp Pro Ala               | Val Thr Leu Ala<br>220                        | His Glu                |
| CTG ATC CAC GCC G<br>Leu Ile His Ala G<br>225 | 230                                   | eu Tyr Gly               | Ile Ala Ile Asn<br>235                        | Pro Asn<br>240         |
|   | 45                                    | sn Ala Tyr<br>250        | Tyr Glu Met Ser                               | Gly Leu<br>255         |
| GAA GTA AGC TTC G<br>Glu Val Ser Phe G<br>260 | iu Giu Leu Ai                         | rg Thr Phe<br>265        | Gly Gly His Asp<br>270                        | Ala Lys                |
| TTT ATC GAC AGC T<br>Phe Ile Asp Ser L<br>275 | eu Gin Giu As<br>28                   | sn Glu Phe<br>30         | Arg Leu Tyr Tyr<br>285                        | Tyr Asn                |
| AAG TTT AAA GAT A<br>Lys Phe Lys Asp I<br>290 | TT GCA AGT AC<br>le Ala Ser Th<br>295 | CA CTG AAC<br>ir Leu Asn | AAG GCT AAG TCC .<br>Lys Ala Lys Ser .<br>300 | ATT GTG 912<br>Ile Val |

- 62 -

| GGT<br>Gly<br>305 | Thr               | ACT               | GCT<br>Ala        | TCA<br>Ser        | TTA<br>Leu<br>310 | Gln               | TAT               | ATG<br>Met        | AAA<br>Lys        | AAT<br>Asn<br>315 | Val               | TTT<br>Phe        | AAA<br>Lys        | GAG<br>Glu        | AAA<br>Lys<br>320 | 2   | 960 |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-----|-----|
| TAT<br>Tyr        | CTC               | CTA<br>Leu        | TCT<br>Ser        | GAA<br>Glu<br>325 | Asp               | ACA<br>Thr        | TCT               | GGA<br>Gly        | Lys<br>330        | Phe               | TCG               | GTA<br>Val        | GAT<br>Asp        | AAA<br>Lys<br>335 | TTA<br>Leu        | 10  | 800 |
| AAA<br>Lys        | TTT<br>Phe        | GAT<br>Asp        | AAG<br>Lys<br>340 | Leu               | TAC               | AAA<br>Lys        | ATG<br>Met        | TTA<br>Leu<br>345 | Thr               | GAG<br>Glu        | ATT               | TAC<br>Tyr        | ACA<br>Thr<br>350 | GAG<br>Glu        | GAT<br>Asp        | 10  | )56 |
| AAT<br>Asn        | TTT               | GTT<br>Val<br>355 | Lys               | TTT<br>Phe        | TTT               | AAA<br>Lys        | GTA<br>Val<br>360 | CTT               | AAC<br>Asn        | AGA<br>Arg        | AAA<br>Lys        | ACA<br>Thr<br>365 | TAT<br>Tyr        | TTG<br>Leu        | AAT<br>Asn        | 11  | .04 |
| TTT<br>Phe        | GAT<br>Asp<br>370 | AAA<br>Lys        | GCC<br>Ala        | GTA<br>Val        | TTT<br>Phe        | AAG<br>Lys<br>375 | ATA<br>Ile        | AAT<br>Asn        | ATA<br>Ile        | GTA<br>Val        | CCT<br>Pro<br>380 | AAG<br>Lys        | GTA<br>Val        | AAT<br>Asn        | TAC<br>Tyr        | 11  | .52 |
| ACA<br>Thr<br>385 | ATA<br>Ile        | TAT<br>Tyr        | GAT<br>Asp        | GGA<br>Gly        | TTT<br>Phe<br>390 | AAT<br>Asn        | TTA<br>Leu        | AGA<br>Arg        | AAT<br>Asn        | ACA<br>Thr<br>395 | AAT<br>Asn        | TTA<br>Leu        | GCA<br>Ala        | Ala               | AAC<br>Asn<br>400 | 12  | 00  |
| TTT<br>Phe        | AAT<br>Asn        | GGT<br>Gly        | CAA<br>Gln        | AAT<br>Asn<br>405 | ACA<br>Thr        | GAA<br>Glu        | ATT<br>Ile        | AAT<br>Asn        | AAT<br>Asn<br>410 | ATG<br>Met        | AAT<br>Asn        | TTT<br>Phe        | ACT<br>Thr        | AAA<br>Lys<br>415 | CTA<br>Leu        | 12  | 48  |
| AAA<br>Lys        | TAA<br>naA        | TTT               | ACT<br>Thr<br>420 | GGA<br>Gly        | TTG<br>Leu        | TTT<br>Phe        | GAA<br>Glu        | TTT<br>Phe<br>425 | TAT<br>Tyr        | AAG<br>Lys        | TTG<br>Leu        | CTA<br>Leu        | TGT<br>Cys<br>430 | GTA<br>Val        | AGA<br>Arg        | 12  | 96  |
| GGG               | ATA<br>Ile        | ATA<br>Ile<br>435 | ACT<br>Thr        | TCT<br>Ser        | AAA<br>Lys        | ACT               | AAA<br>Lys<br>440 | TCA<br>Ser        | TTA<br>Leu        | GAT<br>Asp        | AAA<br>Lys        | GGA<br>Gly<br>445 | TAC               | AAT<br>Asn        | AAG<br>Lys        | 13  | 44  |
| ATC<br>Ile        | GAA<br>Glu<br>450 | GGT<br>Gly        | CGT<br>Arg        | TGC<br>Cys        | GAT<br>Asp        | GGG<br>Gly<br>455 | GCA<br>Ala        | TTA<br>Leu        | AAT<br>Asn        | GAT<br>Asp        | TTA<br>Leu<br>460 | TGT<br>Cys        | ATC<br>Ile        | AAA<br>Lys        | GTT<br>Val        | 13  | 92  |
| AAT<br>Asn<br>465 | AAT<br>Asn        | TGG<br>Trp        | GAC<br>Asp        | TTG<br>Leu        | TTT<br>Phe<br>470 | TTT<br>Phe        | AGT<br>Ser        | CCT               | TCA<br>Ser        | GAA<br>Glu<br>475 | GAT<br>Asp        | AAT<br>Asn        | TTT<br>Phe        | Thr               | AAT<br>Asn<br>480 | 144 | 40  |
| GAT<br>Asp        | CTA<br>Leu        | AAT<br>Asn        | AAA<br>Lys        | GGA<br>Gly<br>485 | GAA<br>Glu        | GAA<br>Glu        | ATT<br>Ile        | ACA<br>Thr        | TCT<br>Ser<br>490 | GAT<br>Asp        | ACT<br>Thr        | AAT<br>Asn        | ATA<br>Ile        | GAA<br>Glu<br>495 | GCA<br>Ala        | 140 | 88  |
| GCA<br>Ala        | GAA<br>Glu        | GAA<br>Glu        | AAT<br>Asn<br>500 | ATT<br>Ile        | AGT<br>Ser        | TTA<br>Leu        | GAT<br>Asp        | TTA<br>Leu<br>505 | ATA<br>Ile        | CAA<br>Gln        | CAA<br>Gln        | TAT<br>Tyr        | TAT<br>Tyr<br>510 | TTA<br>Leu        | ACC<br>Thr        | 153 | 36  |
| TTT<br>Phe        | AAT<br>Asn        | TTT<br>Phe<br>515 | GAT<br>Asp        | TAA<br>Asn        | GAA<br>Glu        | CCT<br>Pro        | GAA<br>Glu<br>520 | AAT<br>Asn        | ATT               | TCA<br>Ser        | ATA<br>Ile        | GAA<br>Glu<br>525 | AAT<br>Asn        | CTT<br>Leu        | TCA<br>Ser        | 158 | 84  |
| AGT<br>Ser        | GAC<br>Asp<br>530 | ATT<br>Ile        | ATA<br>Ile        | GGC<br>Gly        | CAA<br>Gln        | TTA<br>Leu<br>535 | GAA<br>Glu        | CTT<br>Leu        | ATG<br>Met        | CCT<br>Pro        | AAT<br>Asn<br>540 | ATA<br>Ile        | GAA<br>Glu        | AGA<br>Arg        | TTT<br>Phe        | 163 | 32  |
| CCT<br>Pro<br>545 | AAT<br>Asn        | GGA<br>Gly        | AAA<br>Lys        | AAG<br>Lys        | TAT<br>Tyr<br>550 | GAG<br>Glu        | TTA<br>Leu        | GAT<br>Asp        | Lys               | TAT<br>Tyr<br>555 | ACT<br>Thr        | ATG<br>Met        | TTC<br>Phe        | CAT<br>His        | TAT<br>Tyr<br>560 | 168 | 80  |
| CTT<br>Leu        | CGT<br>Arg        | GCT<br>Ala        | CAA<br>Gln        | GAA<br>Glu<br>565 | TTT<br>Phe        | GAA<br>Glu        | CAT<br>His        | GGT<br>Gly        | AAA<br>Lys<br>570 | TCT<br>Ser        | AGG<br>Arg        | ATT<br>Ile        | GCT<br>Ala        | TTA<br>Leu<br>575 | ACA<br>Thr        | 172 | 28  |

|                   |                    |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                     |                   | •    |
|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---------------------|-------------------|------|
| AAT<br>Asn        | TCT<br>Ser         | GTT<br>Val        | AAC<br>Asn<br>580 | GAA<br>Glu        | GCA<br>Ala        | TTA<br>Leu        | TTA<br>Leu        | AAT<br>Asn<br>585 | CCT<br>Pro        | AGT<br>Ser        | CGT<br>Arg        | GTT<br>Val        | TAT<br>Tyr<br>590 | ACA<br>Thr          | TTT<br>Phe        | 1776 |
| TTT<br>Phe        | TCT<br>Ser         | TCA<br>Ser<br>595 | GAC<br>Asp        | TAT<br>Tyr        | GTA<br>Val        | AAG<br>Lys        | AAA<br>Lys<br>600 | GTT<br>Val        | AAT<br>Asn        | AAA<br>Lys        | GCT<br>Ala        | ACG<br>Thr<br>605 | GAG<br>Glu        | GCA<br>Ala          | GCT<br>Ala        | 1824 |
| ATG<br>Met        | TTT<br>Phe<br>610  | TTA<br>Leu        | GGC<br>Gly        | TGG<br>Trp        | GTA<br>Val        | GAA<br>Glu<br>615 | CAA<br>Gln        | TTA<br>Leu        | GTA<br>Val        | TAT<br>Tyr        | GAT<br>Asp<br>620 | Phe               | ACC<br>Thr        | GAT<br>Asp          | GAA<br>Glu        | 1872 |
| ACT<br>Thr<br>625 | AGC<br>Ser         | GAA<br>Glu        | GTA<br>Val        | AGT<br>Ser        | ACT<br>Thr<br>630 | ACG<br>Thr        | GAT<br>Asp        | AAA<br>Lys        | Ile               | GCG<br>Ala<br>635 | GAT<br>Asp        | ATA<br>Ile        | ACT<br>Thr        | ATA<br>Ile          | ATT<br>Ile<br>640 | 1920 |
| ATT               | CCA<br>Pro         | TAT<br>Tyr        | ATA<br>Ile        | GGA<br>Gly<br>645 | CCT<br>Pro        | GCT<br>Ala        | TTA<br>Leu        | AAT<br>Asn        | ATA<br>Ile<br>650 | GGT<br>Gly        | AAT<br>Asn        | ATG<br>Met        | TTA<br>Leu        | TAT<br>Tyr<br>655   | AAA<br>Lys        | 1968 |
| GAT<br>Asp        | GAT<br><u>A</u> sp | TTT<br>Phe        | GTA<br>Val<br>660 | GGT<br>Gly        | GCT<br>Ala        | TTA<br>Leu        | ATA<br>I-le       | TTT<br>Phe<br>665 | TCA<br>Ser        | GGA<br>Gly        | GCT<br>Ala        | GTT<br>Val        | ATT<br>Ile<br>670 | Ç <u>T</u> G<br>Leu | TTA<br>Leu        | 2016 |
| GAA<br>Glu        | TTT<br>Phe         | ATA<br>Ile<br>675 | CCA<br>Pro        | GAG<br>Glu        | ATT<br>Ile        | GCA<br>Ala        | ATA<br>Ile<br>680 | CCT<br>Pro        | GTA<br>Val        | TTA<br>Leu        | GGT<br>Gly        | ACT<br>Thr<br>685 | TTT<br>Phe        | GCA<br>Ala          | CTT<br>Leu        | 2064 |
| GTA<br>Val        | TCA<br>Ser<br>690  | TAT               | ATT               | GCG<br>Ala        | AAT<br>Asn        | AAG<br>Lys<br>695 | GTT<br>Val        | CTA<br>Leu        | ACC<br>Thr        | GTT<br>Val        | CAA<br>Gln<br>700 | ACA<br>Thr        | ATA<br>Ile        | GAT<br>Asp          | AAT<br>Asn        | 2112 |
| GCT<br>Ala<br>705 | TTA<br>Leu         | agt<br>Ser        | AAA<br>Lys        | AGA<br>Arg        | AAT<br>Asn<br>710 | GAA<br>Glu        | AAA<br>Lys        | TGG<br>Trp        | GAT<br>Asp        | GAG<br>Glu<br>715 | GTC<br>Val        | TAT<br>Tyr        | AAA<br>Lys        | TAT<br>Tyr          | ATA<br>Ile<br>720 | 2160 |
| GTA<br>Val        | ACA<br>Thr         | AAT<br>Asn        | TGG<br>Trp        | TTA<br>Leu<br>725 | GCA<br>Ala        | AAG<br>Lys        | GTT<br>Val        | TAA<br>Asn        | ACA<br>Thr<br>730 | CAG<br>Gln        | ATT<br>Ile        | GAT<br>Asp        | CTA<br>Leu        | ATA<br>Ile<br>735   | AGA<br>Arg        | 2208 |
| AAA<br>Lys        | aaa<br>Lys         | ATG<br>Met        | AAA<br>Lys<br>740 | GAA<br>Glu        | GCT<br>Ala        | TTA<br>Leu        | GAA<br>Glu        | AAT<br>Asn<br>745 | CAA<br>Gln        | GCA<br>Ala        | GAA<br>Glu        | GCA<br>Ala        | ACA<br>Thr<br>750 | AAG<br>Lys          | GCT<br>Ala        | 2256 |
| ATA<br>Ile        | ATA<br>Ile         | AAC<br>Asn<br>755 | TAT<br>Tyr        | CAG-<br>Gln       | TAT.<br>Tyr       | AAT<br>Asn        | CAA<br>Gln<br>760 | TAT<br>Tyr        | ACT<br>Thr        | GAG<br>Glu        | GAA<br>Glu        | GAG<br>Glu<br>765 | aaa<br>Lys        | AAT<br>Asn          | AAT<br>Asn        | 2304 |
| ATT<br>Ile        | AAT<br>Asn<br>770  | TTT<br>Phe        | AAT<br>Asn        | ATT<br>Ile        | GAT<br>Asp        | GAT<br>Asp<br>775 | TTA<br>Leu        | AGT<br>Ser        | TCG<br>Ser        | AAA<br>Lys        | CTT<br>Leu<br>780 | AAT<br>Asn        | GAG<br>Glu        | TCT                 | ATA<br>Ile        | 2352 |
| AAT<br>Asn<br>785 | AAA<br>Lys         | GCT<br>Ala        | ATG<br>Met        | ATT<br>Ile        | AAT<br>Asn<br>790 | ATA<br>Ile        | AAT<br>Asn        | AAA<br>Lys        | TTT<br>Phe        | TTG<br>Leu<br>795 | AAT<br>Asn        | CAA<br>Gln        | TGC<br>Cys        | TCT<br>Ser          | GTT<br>Val<br>800 | 2400 |
| TCA<br>Ser        | TAT<br>Tyr         | TTA<br>Leu        | ATG<br>Met        | TAA<br>Asn<br>805 | TCT<br>Ser        | ATG<br>Met        | ATC<br>Ile        | CCT<br>Pro        | TAT<br>Tyr<br>810 | GGT<br>Gly        | GTT<br>Val        | AAA<br>Lys        | CGG<br>Arg        | TTA<br>Leu<br>815   | GAA<br>Glu        | 2448 |
| GAT<br>Asp        | TTT<br>Phe         | GAT<br>Asp        | GCT<br>Ala<br>820 | AGT<br>Ser        | CTT<br>Leu        | AAA<br>Lys        | GAT<br>Asp        | GCA<br>Ala<br>825 | TTA<br>Leu        | TTA<br>Leu        | AAG<br>Lys        | TAT<br>Tyr        | ATA<br>Ile<br>830 | TAT<br>Tyr          | GAT<br>Asp        | 2496 |
| AAT<br>Asn        | AGA<br>Arg         | GGA<br>Gly<br>835 | ACT<br>Thr        | TTA<br>Leu        | ATT<br>Ile        | GGT<br>Gly        | CAA<br>Gln<br>840 | GTA<br>Val        | GAT<br>Asp        | AGA<br>Arg        | TTA<br>Leu        | AAA<br>Lys<br>845 | GAT<br>Asp        | AAA<br>Lys          | GTT<br>Val        | 2544 |

2592

2637

|                   |                   |            |               |                                 |                   |                   |             |            |            | - 64               |                   |            |            |            |            |     |   |
|-------------------|-------------------|------------|---------------|---------------------------------|-------------------|-------------------|-------------|------------|------------|--------------------|-------------------|------------|------------|------------|------------|-----|---|
| AAT<br>Asn        | AAT<br>Asn<br>850 | ACA<br>Thr | CTT           | AGT<br>Ser                      | ACA<br>Thr        | GAT<br>Asp<br>855 | ATA<br>Ile  | CCT<br>Pro | TTT<br>Phe | CAG<br>Gln         | CTT<br>Leu<br>860 | TCC        | AAA<br>Lys | TAC<br>Tyr | GTA<br>Val |     |   |
| GAT<br>Asp<br>865 | AAT<br>Asn        | CAA<br>Gln | AGA<br>Arg    | TTA<br>Leu                      | TTA<br>Leu<br>870 | TCT<br>Ser        | ACA         | TTT        | ACT<br>Thr | GAA<br>Glu<br>875  | TAT               | ATT        | AAG<br>Lys | TAA<br>*   |            |     |   |
| (2)               | INF               | ORMA!      | rion          | FOR                             | SEQ               | ID 1              | NO: 3       | 12:        |            |                    |                   |            |            |            | •          | •   |   |
|                   |                   | ()<br>()   | A) Li<br>B) T | ENCE<br>ENGTI<br>YPE :<br>OPOLO | H: 8'<br>amin     | 79 a              | mino<br>cid |            |            |                    |                   |            |            | ٠          |            |     |   |
|                   |                   |            |               | CE DI                           |                   |                   |             | SEQ :      | ID NO      | ): 12              | 2:                | •          |            |            | •          |     |   |
| Met<br>1          | Gln               | Phe        | Val           | Asn<br>5                        | Lys               | Gln               | Phe         | Asn        | Tyr<br>10  | Lys                | Asp               | Pro        | Val        | Asn<br>15  | Gly        | . : |   |
| Val               | Asp               | Ile        | Ala<br>20     | Tyr                             | Ile               | Lys               | Ile         | Pro<br>25  | Asn        | Ala                | Gly               | Gln        | Met<br>.30 | Gln        | Pro        |     |   |
| Val               | Lys               | Ala<br>35  | Phe           | Lys                             | Ile               | His               | Asn<br>4.0  | Lys        | Ile        | Trp                | Val               | Ile<br>45  | Pro        | Glu        | Arg        | )   |   |
| qaA               | Thr<br>50         | Phe        | Thr           | Asn                             | Pro               | Glu<br>55         | Glu         | Gly        | Asp        | Leu                | Asn<br>60         | Pro        | Pro        | Pro        | Glu        |     |   |
| Ala<br>65         | Lys               | Gln        | Val           | Pro                             | Val<br>70         | Ser               | Tyr         | Tyr        | Asp        | Ser<br>75          | Thr               | Tyr        | Leu        | Ser        | Thr<br>80  |     |   |
| Asp               | Asn               | Glu        | Lys           | Asp<br>85                       | Asn               | Tyr               | Leu         | Lys        | Gly<br>90  | Val                | Thr               | Lys        | Leu        | Phe<br>95  | Glu        |     |   |
| Arg               | Ile               | Tyr        | Ser<br>100    | Thr                             | Asp               | Leu               | Gly         | Arg<br>105 |            | Leu                | Leu               | Thr        | Ser<br>110 | Ile        | Val        | •   |   |
| Arg               | Gly               | Ile<br>115 | Pro           | Phe                             | Trp               | Gly               | Gly<br>120  | Ser        | Thr        | Ile                |                   | Thr<br>125 | Glu        | Leu        | Lys        | . • |   |
| Val               | Ile<br>130        | Asp        | Thr           | Asn                             | Cys               | Ile<br>135        | Asn         | Val        | Ile        | Gln                | Pro<br>140        | qaA        | Gly        | Ser        | Tyr        | ÷   |   |
| Arg<br>145        | Ser               | Glu        | Glu           | Leu                             | Asn<br>150        | Leu               | Val         | Ile        | Ile        | Gly<br><b>15</b> 5 | Pro               | Ser        | Ala        | Asp        | Ile<br>160 |     |   |
| Ile               | Gln               | Phe        | Glu           | Сув<br>165                      | Lys               | Ser               | Phe         | Gly        | His<br>170 | Glu                | Val               | Leu        | Asn        | Leu<br>175 | Thr        |     | , |
| Arg               | Asn               | Gly        | Tyr<br>180    | Gly                             | Ser               | Thr               | Gln         | Tyr<br>185 | Ile        | Arg                | Phe               | Ser        | Pro<br>190 | Asp        | Phe        |     |   |

Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu 195 200 205

Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu 210 215 220

Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn

Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu 250 255

- 65 -

Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys 265 Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr 370 380 Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn 395 Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu 410 Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys Ile Glu Gly Arg Cys Asp Gly Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn 470 475 Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala 490 Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr 505 Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser 520 Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr 545 550 550 Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe 585 Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala

Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile 630 635 Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys 645 Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu 680 Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile 710 Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg 725 Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val 855 Asp Asn Gln Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys \* **B70** 

## (2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2862 base pairs (B) TYPE: nucleic acid

  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION:1..2862
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

|                   | 1                 |                    |                    | <b>V</b> 41        | AAC .<br>Asn<br>5     | uys           | GIN                 | rne                | ASŊ              | Ty           | r Ly<br>O         | rs A.             | sp P              | ro '               | Val            | <b>As</b> : | n Gly             | ,   |
|-------------------|-------------------|--------------------|--------------------|--------------------|-----------------------|---------------|---------------------|--------------------|------------------|--------------|-------------------|-------------------|-------------------|--------------------|----------------|-------------|-------------------|-----|
|                   |                   |                    |                    | 20                 | TAC I                 | 116           | uys .               | iie .              | 25               | ASI          | n Al              | a G               | ly G              | ln i               | 1et<br>30      | Glr         | Pro               |     |
| G"<br>Vá          | rg A<br>al L      | AG G<br>ys A       | CT (<br>la )<br>35 | TTC .<br>Phe       | AAG 1<br>Lys 1        | ATT (<br>le i | CAT /               | ASD 1<br>40        | AAA<br>Lys       | ATO          | TG<br>Tr          | G GT<br>p Va      | al I              | TT (<br>le I<br>45 | CCG<br>Pro     | GAA<br>Glu  | CGC<br>Arg        | 144 |
| GJ<br>As          | ·2 -              | CA T<br>hr P<br>50 | TT ;               | ACG A              | AAC C<br>Asn E        | ro G          | AA .G<br>lu G<br>55 | AA (<br>lu (       | GA<br>Sly        | GAC<br>Asp   | TTO<br>Let        | u As              | AC CO<br>in Pi    | CG C               | CG<br>To       | CCG<br>Pro  | GAA<br>Glu        | 192 |
| 6                 | 5                 | , 5                |                    | , G. T. 1          | CCA G                 | 70            | er r                | yr 1               | yr               | Asp          | Se <sub>1</sub>   | Th                | r T               | /r L               | eu :           | Ser         | Thr<br>80         | 240 |
|                   | <b>P</b>          |                    |                    | , y 5, F           | SAT A<br>Asp A<br>85  | 211 I         | Ar P                | eu L               | ys               | 90<br>GTA    | Val               | . Th              | r Ly              | 's L               | eu !           | Phe<br>95   | Glu               | 288 |
| CG                | T AT<br>g Il      | T TI<br>e Ty       | , ,                | CC A<br>er T<br>00 | CT G.                 | AC C          | TG G                | TA W               | GT<br>rg i<br>05 | ATG<br>Met   | CTG<br>Leu        | CT<br>Le          | G AC<br>u Th      | r Se               | CA A<br>er J   | ATC<br>[le  | GTC<br>Val        | 336 |
| ***               | 5 01              | 11                 | .5                 | LO P               | TT TO                 | ip G.         | 12                  | 20<br>20           | er :             | Thr          | Ile               | Ası               | 7 Th              | r G]<br>5          | lu I           | eu          | Lys               | 384 |
|                   | 13                | 0                  | , p 11             | ** ~               | AC TO                 | 13            | .e As               | in va              | al j             | rie          | Gin               | Pro               | As <sub>i</sub>   | p Gl               | y S            | er          | Tyr               | 432 |
| AGA<br>Arg        | ,                 | r GA<br>r Gl       | A G/<br>u Gl       | AA C'              | TT AA<br>eu As<br>15  | n re          | C GI<br>u Va        | 'A A'I             | C A              | TC<br>le     | GGG<br>Gly<br>155 | CCC               | TC(               | C GC<br>c Al       | G G<br>a A     | sp          | ATT<br>Ile<br>160 | 480 |
|                   |                   |                    |                    | 16                 | -                     | 5 <b>5</b> E  | r Pn                | e GI               | у н<br>1         | 70           | Glu               | Val               | Let               | l As               | n Lo           | eu '<br>75  | Thr               | 528 |
| Arg               | AAC               | GG<br>Gl           | Т ТА<br>у Ту<br>18 | - 91               | Y Se                  | T AC          | T CA                | G TA<br>n Ty<br>18 | L I              | TT  <br>le / | CGT<br>Arg        | TTC<br>Phe        | AGC<br>Ser        | Pro                | D As           | AC :        | Phe               | 576 |
| ACG<br>Thr        | TTC               | GG;<br>Gl;<br>199  | F 11               | C GA<br>e Gl       | G GA                  | G AG          | C CTO               | T GT               | G G'<br>u V      | TT (         | GAT<br>Asp        | ACC<br>Thr        | AAC<br>Asn<br>205 | Pro                | G CT           | G I         | rrg<br>Leu        | 624 |
| GGT<br>Gly        | GCA<br>Ala<br>210 | O L y              | Ly                 | G TT<br>s Ph       | C GCA<br>e Ala        | A ACT         | Ası                 | CC:                | A GO             | CG (         | Val               | ACC<br>Thr<br>220 | CTG<br>Leu        | GCA<br>Ala         | CA<br>Hi       | .C G        | AG<br>lu          | 672 |
| CTG<br>Leu<br>225 | ATC<br>Ile        | CAC                | GC0<br>Ala         | C GG<br>a Gl       | T CAT<br>Y His<br>230 | AFC           | CTC<br>Leu          | TA:                | r Gi             | ry 1         | ATT (<br>le )     | GCG<br>Ala        | ATT<br>Ile        | AAC<br>Asn         | CC<br>Pr       | o A         | AC<br>sn<br>40    | 720 |
| CGC<br>Arg        | GTG<br>Val        | TTC                | Lys                | GT:<br>Va:<br>24!  | r aac<br>l asn        | ACC           | AAC<br>Asn          | GCC<br>Ala         | TA<br>Ty<br>25   | T            | AC (              | GAG<br>Glu        | ATG<br>Met        | AGT<br>Ser         | GG<br>G1<br>25 | y L         | TA<br>eu          | 768 |
| GAA<br>Glu        | GTA<br>Val        | AGC<br>Ser         | Phe<br>260         | GIL                | GAA<br>Glu            | CTG           | CGC                 | ACG<br>Thr<br>265  | Ph               | 'C G<br>le G | GT (              | GGC<br>Gly        | CAT<br>His        | GAT<br>Asp<br>270  | GC(            | G A.<br>a L | AG<br>Ys          | 816 |

| TTT<br>Phe        | ATC<br>Ile        | GAC<br>Asp<br>275 | AGC<br>Ser  | TTG<br>Leu        | CAG<br>Gln        | GAG<br>Glu        | AAC<br>Asn<br>280 | Glu        | TTC<br>Phe        | CGT<br>Arg        | CTG<br>Leu        | TAC<br>Tyr<br>285 | TAC        | TAC               | AAC<br>Asn        |     | 864  |
|-------------------|-------------------|-------------------|-------------|-------------------|-------------------|-------------------|-------------------|------------|-------------------|-------------------|-------------------|-------------------|------------|-------------------|-------------------|-----|------|
| AAG<br>Lys        | TTT<br>Phe<br>290 | AAA<br>Lys        | GAT<br>Asp  | ATT               | GCA<br>Ala        | AGT<br>Ser<br>295 | ACA<br>Thr        | CTG<br>Leu | AAC<br>Asn        | AAG<br>Lys        | GCT<br>Ala<br>300 | AAG<br>Lys        | TCC<br>Ser | ATT<br>Ile        | GTG<br>Val        |     | 912  |
| GGT<br>Gly<br>305 | ACC<br>Thr        | ACT               | 'GCT<br>Ala | TCA<br>Ser        | Leu<br>310        | CAG               | TAT<br>Tyr        | ATG<br>Met | AAA<br>Lys        | AAT<br>Asn<br>315 | GTT<br>Val        | TTT<br>Phe        | AAA<br>Lys | GAG<br>Glu        | AAA<br>Lys<br>320 |     | 960  |
| TAT<br>Tyr        | CTC<br>Leu        | CTA<br>Leu        | TCT<br>Ser  | GAA<br>Glu<br>325 | GAT<br>Asp        | ACA<br>Thr        | TCT<br>Ser        | GGA<br>Gly | AAA<br>Lys<br>330 | TTT<br>Phe        | TCG<br>Ser        | GTA<br>Val        | GAT<br>Asp | AAA<br>Lys<br>335 | TTÀ<br>Leu        |     | 1008 |
| Lys               | · Phe             | Asp               | Lys<br>340  | Leu               | Tyr               | AAA<br>Lys        | Met               | Leu<br>345 | Thr               | Glu               | Ile               | Tyr               | Thr<br>350 | Glu               | Asp               |     | 1056 |
| Asn               | Phe               | Val<br>355        | Lys         | Phe               | Phe               | AAA<br>Lys        | Val<br>360        | Leu        | Asn               | Arg               | Lys               | Thr<br>365        | Tyr        | Leu               | Asn               |     | 1104 |
| Phe               | Asp<br>370        | Lys               | Ala         | Val               | Phe               | AAG<br>Lys<br>375 | Ile               | Asn        | Ile               | Val               | Pro<br>380        | Lys               | Val        | Asn               | Tyr               |     | 1152 |
| Thr<br>385        | Ile               | Tyr               | Asp         | Gly               | Phe<br>390        | AAT<br>Asn        | Leu               | Arg        | Asn               | Thr<br>395        | Asn               | Leu               | Ala        | Ala               | Asn<br>400        |     | 1200 |
| Phe               | Asn               | Gly               | Gln         | Asn<br>405        | Thr               | GAA<br>Glu        | Ile               | Asn        | Asn<br>410        | Met               | Asn               | Phe               | Thr        | Lys<br>415        | Leu.              |     | 1248 |
| Lys               | Asn               | Phe               | Thr<br>420  | Gly               | Leu               |                   | Glu               | Phe<br>425 | Tyr               | Lys               | Leu               | Leu               | Cys<br>430 | Val               | Arg.              |     | 1296 |
| GGG<br>Gly        | ATA<br>Ile        | ATA<br>Ile<br>435 | ACT         | TCT               | AAA<br>Lys        | ACT<br>Thr        | AAA<br>Lys<br>440 | TCA<br>Ser | TTA<br>Leu        | GAT<br>Asp        | AAA<br>Lys        | GGA<br>Gly<br>445 | TAC<br>Tyr | AAT<br>Asn        | AAG<br>Lys        |     | 1344 |
| Ile               | Glu<br>450        | Gly               | Arg         | Cys               | Asp               | GGG<br>Gly<br>455 | Ala               | Leu        | Asn               | Asp               | Leu<br>460        | Cys               | Ile        | Lys               | Val               |     | 1392 |
| AAT<br>Asn<br>465 | AAT<br>Asn        | TGG<br>Trp        | GAC<br>Asp  | TTG<br>Leu        | TTT<br>Phe<br>470 | TTT<br>Phe        | AGT<br>Ser        | CCT<br>Pro | TCA<br>Ser        | GAA<br>Glu<br>475 | GAT<br>Asp        | AAT<br>Asn        | TTT<br>Phe | ACT<br>Thr        | AAT<br>Asn<br>480 | . • | 1440 |
| GAT<br>Asp        | CTA<br>Leu        | AAT<br>Asn        | AAA<br>Lys  | GGA<br>Gly<br>485 | GAA<br>Glu        | GAA<br>Glu        | ATT               | ACA<br>Thr | TCT<br>Ser<br>490 | GAT<br>Asp        | ACT<br>Thr        | AAT<br>Asn        | ATA<br>Ile | GAA<br>Glu<br>495 | GCA<br>Ala        |     | 1488 |
|                   |                   |                   |             |                   |                   | TTA<br>Leu        |                   |            |                   |                   |                   |                   |            |                   |                   | •   | 1536 |
| TTT Phe           | AAT<br>Asn        | TTT<br>Phe<br>515 | GAT<br>Asp  | AAT<br>Asn        | GAA<br>Glu        | CCT<br>Pro        | GAA<br>Glu<br>520 | AAT<br>Asn | ATT<br>Ile        | TCA<br>Ser        | ATA<br>Ile        | GAA<br>Glu<br>525 | AAT<br>Asn | CTT<br>Leu        | TCA<br>Ser        |     | 1584 |
| AGT<br>Ser        | GAC<br>Asp<br>530 | ATT<br>Ile        | ATA<br>Ile  | GGC<br>Gly        | CAA<br>Gln        | TTA<br>Leu<br>535 | GAA<br>Glu        | CTT<br>Leu | ATG<br>Met        | CCT<br>Pro        | AAT<br>Asn<br>540 | ATA<br>Ile        | GAA<br>Glu | AGA<br>Arg        | TTT<br>Phe        | -   | 1632 |

| •                 |                    |              |                      |                      |                       |                       |                       |                      |                      |                     |                    |                    |                   |                    |                |                   |      |
|-------------------|--------------------|--------------|----------------------|----------------------|-----------------------|-----------------------|-----------------------|----------------------|----------------------|---------------------|--------------------|--------------------|-------------------|--------------------|----------------|-------------------|------|
| 5                 | 15                 | J U          | -y -                 | Jy S L               | _                     | 50                    | TU L                  | eu A                 | ab ri                | /S Ty<br>55         | yr T)<br>55        | nr Me              | et P              | he H               | is             | Tyr<br>560        | 1680 |
| C.                | rr co<br>eu Ai     | GT G         | CT C<br>la G         | TII C                | SAA T<br>Slu P<br>665 | TT G<br>he G          | AA CI<br>lu Hi        | AT GO                | GT AF<br>Ly Ly<br>57 | 's Se               | er Ar              | GG AT              | TT GO             | la L               | TA<br>eu<br>75 | ACA<br>Thr        | 1728 |
| A.A<br>As         | T TO               | T G          | ar 7                 | AC G<br>sn G<br>80   | AA G                  | CA T                  | TA TI<br>eu Le        | AA AT<br>Lu As<br>S8 | in Pr                | T AG                | ST CO              | ST GI              | TT TA             | r T                | CA<br>hr       | TTT<br>Phe        | 1776 |
| TI                | T TC<br>e Se       |              | CA G.<br>er A.<br>95 | AC T<br>sp T         | AT G                  | TA AJ<br>al Ly        | AG AA<br>/s Ly<br>60  | 's va                | T AA<br>l As         | T AA<br>n Ly        | A GC               | T AC<br>a Th<br>60 | r Gl              | G GC               | CA (           | GCT<br>Ala        | 1824 |
| AT<br>Me          | G TT<br>t Ph<br>61 | C 24         | ra Go<br>eu Gi       | GC TO                | GG G'<br>rp Va        | ra ga<br>al gi<br>61  | .u GI                 | A TT<br>n Le         | A GT.<br>u Va        | A TA                | T GA<br>r As<br>62 | p Ph               | T AC<br>e Th      | C GA               | AT (           | GAA<br>Glu        | 1872 |
| AC<br>Th<br>62    |                    | C GA<br>r Gl | iA G7<br>.u Va       | ra ad<br>al Se       | GT AC<br>er Th<br>63  | T III                 | G GA                  | T AA<br>p Ly         | A AT                 | r GCc<br>Ala<br>63! | a Ası              | T AT.              | A AC<br>e Th      | T AT               | e i            | ATT<br>Ile<br>540 | 1920 |
| AT'               | r cc               | A TA         | T AT                 | CA GC<br>Le G1<br>64 | SA CC<br>Ly Pr        | T GC                  | T TT:<br>a Le         | A AA:<br>u Asi       | T ATA                | e Gly               | r aa:<br>Y Asi     | T ATO              | G TT              | A TA<br>L Ty<br>65 | r I            | lys               | 1968 |
| GAT<br>Asp        | GA:                | r TT<br>> Ph | T GT<br>e'Va<br>66   | T GY                 | ST GC<br>.y Al        | T TT.<br>a Le         | A ATA                 | A TTT                | Ser                  | GGA<br>Gly          | A GCT<br>/ Ala     | r GTT<br>a Val     | T ATT             | e Le               | G I            | TA<br>eu          | 2016 |
| GAA<br>Glu        | TTI<br>Phe         | 110<br>67    | C PI                 | A GA<br>o Gl         | G AT                  | T GC                  | A ATA<br>a Ile<br>680 | , hrc                | GTA<br>Val           | TTA<br>Leu          | GGT<br>Gly         | C ACT              | Phe               | GC/<br>Ala         | A C            | TT<br>eu          | 2064 |
| GTA<br>Val        | TCA<br>Ser<br>690  | - Y 2        | r AT                 | T GC<br>e Al         | G AA'<br>a Ası        | T AAC<br>A Lys<br>695 | s val                 | CTA<br>Leu           | ACC<br>Thr           | GTT<br>Val          | CAA<br>Gln<br>700  | Thr                | ATA<br>Ile        | GAT<br>Asp         | T A            | AT<br>sn          | 2112 |
| GCT<br>Ala<br>705 | TTA<br>Leu         | AG1<br>Sei   | Lys                  | A AG                 | A AA7<br>g Asi<br>710 | r GTO                 | AAA<br>Lys            | TGG<br>Trp           | GAT<br>Asp           | GAG<br>Glu<br>715   | GTC<br>Val         | TAT<br>Tyr         | AAA<br>Lys        | TAT<br>Tyr         | : I            | TA<br>le<br>20    | 2160 |
| <b>V</b> 42       | ****               | ASI          |                      | 729                  | -                     | r rys                 | vai                   | Asn                  | 730                  | Gln                 | Ile                | Asp                | Leu               | 11e<br>735         | Aı             | rg                | 2208 |
| AAA<br>Lys        | AAA<br>Lys         | ATG<br>Met   | Lys<br>740           | GIL                  | A GCT                 | TTA<br>Leu            | GAA<br>Glu            | AAT<br>Asn<br>745    | CAA<br>Gln           | GCA<br>Ala          | GAA<br>Glu         | GCA<br>Ala         | ACA<br>Thr<br>750 | AAG<br>Lys         | GC<br>Al       | T<br>.a           | 2256 |
| 116               | 116                | 755          | IYL                  | GII                  | TAT                   | ASD                   | 760                   | Tyr                  | Thr                  | Glu                 | Glu                | Glu<br>765         | Lys               | Asn                | As             | ın                | 2304 |
| ATT<br>Ile        | AAT<br>Asn<br>770  | TTT<br>Phe   | AAT<br>Asn           | ATT                  | GAT<br>Asp            | GAT<br>Asp<br>775     | TTA<br>Leu            | AGT<br>Ser           | TCG<br>Ser           | AAA<br>Lys          | CTT<br>Leu<br>780  | AAT<br>Asn         | GAG<br>Glu        | TCT<br>Ser         | AT<br>Il       | 'A<br>e           | 2352 |
| AAT<br>Asn<br>785 | AAA<br>Lys         | GCT<br>Ala   | ATG<br>Met           | ATT<br>Ile           | AAT<br>Asn<br>790     | ATA<br>Ile            | AAT<br>Asn            | AAA<br>Lys           | TTT<br>Phe           | TTG<br>Leu<br>795   | AAT<br>Asn         | CAA<br>Gln         | TGC<br>Cys        | TCT<br>Ser         | GT<br>Va<br>80 | 1                 | 2400 |
| rca -<br>Ser      | TAT<br>Tyr         | TTA<br>Leu   | ATG<br>Met           | AAT<br>Asn<br>805    | TCT<br>Ser            | ATG<br>Met            | ATC<br>Ile            | Pro                  | TAT<br>Tyr<br>810    | GGT<br>Gly          | GTT<br>Val         | AAA<br>Lys         | Arg               | TTA<br>Leu<br>815  | GA:            | A<br>u            | 2448 |
|                   |                    |              |                      |                      |                       |                       |                       |                      |                      |                     |                    |                    |                   |                    |                |                   |      |

| •                 |                   |                   |                   |                   |                   |                   |                   |                   | ٠.                |                   |                   |                   |                   |                   |                   |   |      |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---|------|
| GAT<br>Asp        | TTT               | GAT<br>Asp        | GCT<br>Ala<br>820 | AGT<br>Ser        | CTT<br>Leu        | AAA<br>Lys        | GAT<br>Asp        | GCA<br>Ala<br>825 | TTA<br>Leu        | TTA<br>Leu        | AAG<br>Lys        | TAT               | ATA<br>Ile<br>830 | TAT               | GAT<br>Asp        |   | 2496 |
| AAT<br>Asn        | AGA<br>Arg        | GGA<br>Gly<br>835 | ACT<br>Thr        | TTA<br>Leu        | ATT<br>Ile        | GGT<br>Gly        | CAA<br>Gln<br>840 | GTA<br>Val        | GAT<br>Asp        | AGA<br>Arg        | TTA<br>Leu        | AAA<br>Lys<br>845 | GAT<br>Asp        | AAA<br>Lys        | GTT<br>Val        | ٠ | 2544 |
| AAT<br>Asn        | AAT<br>Asn<br>850 | Thr               | CTT<br>Leu        | AGT<br>Ser        | ACA<br>Thr        | GAT<br>Asp<br>855 | ATA<br>Ile        | CCT<br>Pro        | TTT<br>Phe        | CAG<br>Gln        | CTT<br>Leu<br>860 | TCC<br>Ser        | AAA<br>Lys        | TAC<br>Tyr        | GTA<br>Val        |   | 2592 |
| GAT<br>Asp<br>865 | AAT<br>Asn        | CAA<br>Glņ        | AGA<br>Arg        | TTA<br>Leu        | TTA<br>Leu<br>870 | TCT               | ACA<br>Thr        | TTT<br>Phe        | ACT<br>Thr        | GAA<br>Glu<br>875 | TAT<br>Tyr        | ATT               | AAG<br>Lys        | TCT<br>Ser        | AGG<br>Arg<br>880 |   | 2640 |
| CCT<br>Pro        | GGA<br>Gly        | CCG<br>Pro        | GAG<br>Glu        | ACG<br>Thr<br>885 | CTC<br>Leu        | TGC<br>Cys        | GGG<br>Gly        | GCT<br>Ala        | GAG<br>Glu<br>890 | CTG<br>Leu        | GTG<br>Val        | GAT<br>Asp        | GCT<br>Ala        | CTT<br>Leu<br>895 | CAG<br>Gln        |   | 2688 |
| TTC<br>Phe        | GTG<br>Val        | TGT<br>Cys        | GGA<br>Gly<br>900 | GAC               | AGG<br>Arg        | GGC<br>Gly        | TTT               | TAT<br>Tyr<br>905 | TTC<br>Phe        | AAC<br>Asn        | AAG<br>Lys        | CCC<br>Pro        | ACA<br>Thr<br>910 | GGG<br>Gly        | TAT<br>Tyr        |   | 2736 |
| GGC<br>Gly        | TCC<br>Ser        | AGC<br>Ser<br>915 | AGT<br>Ser        | CGG<br>Arg        | AGG<br>Arg        | GCG<br>Ala        | CCT<br>Pro<br>920 | CAG<br>Gln        | ACA<br>Thr        | GGT<br>Gly        | ATC<br>Ile        | GTG<br>Val<br>925 | GAT<br>Asp        | GAG<br>Glu        | TGC<br>Cys        |   | 2784 |
| TGC<br>Cys        | TTC<br>Phe<br>930 | CGG<br>Arg        | AGC<br>Ser        | TGT<br>Cys        | GAT<br>Asp        | CTA<br>Leu<br>935 | AGG<br>Arg        | AGG<br>Arg        | CTG<br>Leu        | GAG<br>Glu        | ATG<br>Met<br>940 | TAT<br>Tyr        | TGC<br>Cys        | GCA<br>Ala        | CCC<br>Pro        | - | 2832 |
| CTC<br>Leu<br>945 | AAG<br>Lys        | CCT<br>Pro        | GCC Ala           | AAG<br>Lys        | TCA<br>Ser<br>950 | GCT<br>Ala        | GAA<br>Glu        | GCT<br>Ala        | TAG               | ·<br>•            |                   |                   |                   |                   |                   |   | 2862 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |   |      |

## (2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 954 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly
1 5 10 15

Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro 20 25 30

Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg

Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Glu 50 55 60

Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr 65 70 75 80

Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu 85 90 95

Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val

- 71 -

Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly, Ser Tyr 135 Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile 150 Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu 200 Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu 215 Leu Ile His Ala Gly His Arg Leu Tyr Gly He Ala Ile Asn Pro Asn 230 Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys 260 Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp 345 Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn 395 Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu 405 410 Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys Ile Glu Gly Arg Cys Asp Gly Ala Leu Asn Asp Leu Cys Ile Lys Val

Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn 475 Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala 490 Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr 500 505 Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe 535 Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr 570 Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu 680 Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile 770 780 Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val 790 Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu

|            |            |                           |                   |                        |   |                         |            |            |            | - 73       | -          |            |            |            |            |
|------------|------------|---------------------------|-------------------|------------------------|---|-------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Asn        | Phe        | Asp                       | Ala<br>820        | Ser                    | Leu                                       | Lys                     | Asp        | Ala<br>825 | Leu        | Leu        | Lys        | Tyr        | Ile<br>830 | Tyr        | Asp        |
|            | Arg        | Gly<br>835                | Thr               | Leu                    | Ile                                       | Gly                     | Gln<br>840 | Val        | Asp        | Arg        | Leu        | Lys<br>845 | Asp        | Lys        | Val        |
| Asn        | Asn<br>850 | Thr                       | Leu               | Ser                    | Thr                                       | Asp<br>855              | Ile        | Pro        | Phe        | Gln        | Leu<br>860 | Ser        | Lys        | Tyr        | Val        |
| Asp<br>865 | Asn        | Gln                       | Arg               | Leu                    | Leu<br>870                                | Ser                     | Thr        | Phe        | Thr        | Glu<br>875 | Tyr        | Ile        | Lys        | Ser        | Arg<br>880 |
| Pro        | Gly        | Pro                       | Glu               | Thr<br>885             | Leu                                       | Cys <sup>.</sup>        | Gly        | Ala        | Glu<br>890 | Leu        | Val        | Asp        | Ala        | Leu<br>895 | Gln        |
| Phe        | Val        | Cys                       | Gly<br>900        | Asp                    | Arg                                       | Gly                     | Phe        | Tyr<br>905 | Phe        | Asn        | Lys        | Pro        | Thr<br>910 | Gly        | Tyr        |
| Gly :      | Ser        | Ser<br>915                | Ser               | Arg                    | Arg                                       | Ala                     | Pro<br>920 | Gln        | Thr        | Gly        | Ile        | Val<br>925 | Asp        | Glu        | Cys        |
| lys !      | Phe<br>930 | Arg                       | Ser               | Cys                    | Asp .                                     | Leu :<br>935            | Arg        | Arg        | Leu        | Glu        | Met<br>940 | Tyr        | Cys        | Ala        | Pro        |
| eu I<br>45 | Lys        | Pro                       | Ala               |                        | Ser 2<br>950                              | Ala                     | Glu .      | Ala        | •          |            |            | ,          |            |            |            |
| 2) 1       | INFO       | RMAT:                     | ION               | FOR :                  | SEQ :                                     | ID N                    | 0: 1       | 5 :        |            |            |            |            |            |            |            |
|            | (i)        | SEQ(<br>(A)<br>(B)<br>(C) | LEI<br>TYI<br>STI | NGTH<br>PE: 1<br>RAND! | ARACT<br>: 272<br>nucle<br>EDNES<br>SY: 1 | 24 ba<br>sic a<br>SS: c | ase pacid  | pairs      | 5          |            |            |            |            |            |            |
| (          | ii)        | MOLE                      | CUL               | TY!                    | PE: D                                     | NA (                    | genc       | mic)       |            |            |            |            |            |            |            |
|            | ix)        | (A)                       |                   | IE/KE                  | Y: C<br>N:1.                              |                         | 4          |            |            | -          |            |            |            |            |            |

| (X1) | SEQUENCE | DESCRIPTION: | SEQ | ID | NO: | 15: |  |
|------|----------|--------------|-----|----|-----|-----|--|
|      |          |              |     |    |     |     |  |

| ATG<br>Met       | CAG<br>Gln       | TTC<br>Phe       | GTG<br>Val       | AAC<br>Asn<br>5  | AAG<br>Lys       | CAG<br>Gln       | TTC<br>Phe       | AAC<br>Asn       | TAT<br>Tyr<br>10 | Lys              | GAC<br>Asp       | CCT              | GTA<br>Val       | AAC<br>Asn<br>15 | GGT              | 48  |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|-----|
| GTT<br>Val       | GAC<br>Asp       | ATT              | GCC<br>Ala<br>20 | TAC<br>Tyr       | ATC              | AAA<br>Lys       | ATT              | CCA<br>Pro<br>25 | Asn              | GCC<br>Ala       | GGC<br>Gly       | CAG<br>Gln       | ATG<br>Met<br>30 | CAG<br>Gln       | CCG<br>Pro       | 96  |
| GTG<br>Val       | AAG<br>Lys       | GCT<br>Ala<br>35 | TTC<br>Phe       | AAG<br>Lys       | ATT<br>Ile       | CAT<br>His       | AAC<br>Asn<br>40 | AAA<br>Lys       | ATC<br>Ile       | TGG<br>Trp       | GTT<br>Val       | ATT<br>Ile<br>45 | CCG<br>Pro       | GAA<br>Glu       | CGC<br>Arg       | 144 |
| GAT<br>Asp       | ACA<br>Thr<br>50 | TTT<br>Phe       | ACG<br>Thr       | AAC<br>Asn       | CCG<br>Pro       | GAA<br>Glu<br>55 | GAA<br>Glu       | GGA<br>Gly       | GAC<br>Asp       | TTG<br>Leu       | AAC<br>Asn<br>60 | CCG<br>Pro       | CCG<br>Pro       | CCG<br>Pro       | GAA<br>Glu       | 192 |
| GCA<br>Ala<br>65 | AAG<br>Lys       | CAG<br>Gln       | GTG<br>Val       | CCA<br>Pro       | GTT<br>Val<br>70 | TCA<br>Ser       | TAC<br>Tyr       | TAC<br>Tyr       | GAT<br>Asp       | TCA<br>Ser<br>75 | ACC<br>Thr       | TAT<br>Tyr       | CTG<br>Leu       | AGC<br>Ser       | ACA<br>Thr<br>80 | 240 |
| GAC<br>Asp       | AAC<br>Asn       | GAG<br>Glu       | AAG<br>Lys       | GAT<br>Asp<br>85 | AAC<br>Asn       | TAC<br>Tyr       | CTG<br>Leu       | AAG<br>Lys       | GGA<br>Gly<br>90 | GTG<br>Val       | ACC<br>Thr       | AAA<br>Lys       | TTA<br>Leu       | TTC<br>Phe<br>95 | GAG<br>Glu       | 288 |

| CC<br>A1   | ST AT              | T TA<br>e Ty       | T TC<br>T Se<br>10 |                   | r Ga<br>r As <sub>l</sub> | C CTO          | G GG(                 | C CG:                   | g met          | G CTO      | G CTO             | G AC                  | C TC<br>r Se | r Il       | C GTC<br>e Val |       | 3,36 |
|------------|--------------------|--------------------|--------------------|-------------------|---------------------------|----------------|-----------------------|-------------------------|----------------|------------|-------------------|-----------------------|--------------|------------|----------------|-------|------|
| CG<br>Ar   | c GG<br>g Gl       | A AT<br>y Il<br>11 | C FI               | A TTT             | r TG(<br>e Tr             | G GGT<br>P Gly | GGC<br>Gly<br>120     | Sei                     | C ACC          | Ile        | GAC<br>S Asp      | 2 ACC<br>2 Thi<br>125 | r Glı        | G TT       | G AAG<br>u Lys |       | 384  |
| GT<br>Va   | T AT<br>1 I1<br>13 |                    | C AC               | r AAC<br>r Asr    | TGC<br>Cys                | E ATT          | : ASI                 | GTC<br>Val              | ATC            | CAA<br>Gln | CCA<br>Pro<br>140 | Asp                   | GGT<br>Gly   | C AGO      | TAC<br>Tyr     |       | 432  |
| 14         | 5                  |                    | -                  | , per             | 150                       | )              | vat                   | TIE                     | lle            | 155        | Pro               | Ser                   | Ala          | Asţ        | C ATT          |       | 480  |
|            | . J.               |                    | . GIU              | 165               | Dys                       | i              | Pne                   | GIÀ                     | 170            | Glu        | Val               | Leu                   | . Asn        | Leu<br>175 |                |       | 528  |
|            | , nat              | . Gij              | 180                | GIY               | Ser                       | inr            | Gin                   | 19r<br>185              | Ile            | Arg        | Phe               | Ser                   | Pro<br>190   | qzA        | TTC Phe        |       | 576  |
|            | · FIIC             | 195                | ) .                | GIU               | GIU                       | Ser            | 200                   | GIU                     | Val            | Asp        | Thr               | Asn<br>205            | Pro          | Leu        | TTG<br>Leu     |       | 624  |
| G.         | 210                | GLY                | гуз                | Pne               | Ala                       | 215            | Asp                   | Pro                     | Ala            | Val        | Thr<br>220        | Leu                   | Ala          | His        |                | 3)) * | 672  |
| 225        | 116                | nis                | nia                | GGT<br>Gly        | 230                       | Arg            | Leu                   | lyr                     | GIÀ            | 11e<br>235 | Ala               | Ile                   | Asn          | Pro        | Asn<br>240     |       | 720  |
| nrg        | val                | FIIE               | r\a                | GTT<br>Val<br>245 | Asn                       | Inr            | Asn                   | Ala                     | Tyr<br>250     | Tyr        | Glu               | Met                   | Ser          | Gly<br>255 | Leu            |       | 768  |
| GIU        | AGI                | Ser                | 260                | GAG<br>Glu        | GIU                       | Leu            | Arg                   | 265                     | Phe            | Gly        | Gly               | His                   | Asp<br>270   | Ala        | Lys            |       | 816  |
| PIIC       | 116                | 275                | ser                | TTG<br>Leu        | ĠIĦ                       | GIU            | 280                   | GIu                     | Phe            | Arg        | Leu               | Tyr<br>285            | Tyr          | Tyr        | Asn            |       | 864  |
| nys        | 290                | Lys                | Asp                | ATT<br>Ile        | Ala                       | Ser<br>295     | Tnr                   | Leu                     | Asn            | Lys        | Ala<br>300        | Lys                   | Ser          | Ile        | Val            |       | 912  |
| 305        | Inr                | inr                | ATA                | TCA<br>Ser        | 110                       | GIn            | Tyr                   | Met.                    | Lys .          | Asn<br>315 | Val               | Phe                   | Lys          | Glu        | Lys<br>320     | w     | 960  |
| TYE        | ren                | Leu                | ser                | GAA<br>Glu<br>325 | Asp                       | Tnr            | Ser (                 | Gly                     | 130            | Phe        | Ser               | Val                   | Asp          | Lys<br>335 | Leu            |       | 1008 |
| гÀ2        | Pne                | Asp                | 140                | TTA<br>Leu        | Tyr                       | Lys            | Met :                 | Leu <sup>-</sup><br>345 | Thr (          | Glu        | Ile '             | Tyr                   | Thr<br>350   | Glu        | Asp            | *     | 1056 |
| AAT<br>Asn | TTT<br>Phe         | GTT<br>Val<br>355  | AAG<br>Lys         | TTT '             | TTT .<br>Phe              | Lys. `         | GTA (<br>Val 1<br>360 | CTT .<br>Leu .          | AAC A<br>Asn A | AGA A      | Lys '             | ACA 'Thr '365         | TAT<br>Tyr   | TTG<br>Leu | AAT<br>Asn     |       | 1104 |

| TTT               | GAT<br>Asp<br>370 | AAA<br>Lys        | GCC<br>Ala        | GTA<br>Val        | TTT<br>Phe        | AAG<br>Lys<br>375 | ATA<br>Ile        | AAT<br>Asn        | ATA<br>Ile        | GTA<br>Val        | CCT<br>Pro<br>380 | AAG<br>Lys        | GTA<br>Val        | AAT<br>Asn        | TAC<br>Tyr        |   | 1152 |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---|------|
| ACA<br>Thr<br>385 | ATA<br>Ile        | TAT<br>Tyr        | GAT<br>Asp        | GGA<br>Gly        | TTT<br>Phe<br>390 | TAA<br>naA        | TTA<br>Leu        | AGA<br>Arg        | AAT<br>Asn        | ACA<br>Thr<br>395 | AAT<br>Asn        | TTA<br>Leu        | GCA<br>Ala        | GCA<br>Ala        | AAC<br>Asn<br>400 |   | 1200 |
| TTT               | AAT<br>Asn        | GGT<br>Gly        | CAA<br>Gln        | AAT<br>Asn<br>405 | ACA<br>Thr        | GAA<br>Glu        | ATT<br>Ile        | AAT<br>Asn        | AAT<br>Asn<br>410 | ATG<br>Met        | AAT<br>Asn        | TTT<br>Phe        | ACT<br>Thr        | AAA<br>Lys<br>415 | CTA<br>Leu        |   | 1248 |
| AAA<br>Lys        | AAT<br>Asn        | TTT               | ACT<br>Thr<br>420 | GGA<br>Gly        | TTG<br>Leu        | TTT<br>Phe        | GAA<br>Glu        | TTT<br>Phe<br>425 | TAT<br>Tyr        | AAG<br>Lys        | TTG<br>Leu        | CTA<br>Leu        | TGT<br>Cys<br>430 | GTA<br>Val        | AGA<br>Arg        |   | 1296 |
| GGG<br>Gly        | ATA<br>Ile        | ATA<br>Ile<br>435 | ACT<br>Thr        | TCT<br>Ser        | AAA<br>Lys        | ACT<br>Thr        | AAA<br>Lys<br>440 | TCA<br>Ser        | TTA<br>Leu        | GAT<br>Asp        | AAA<br>Lys        | GGA<br>Gly<br>445 | TAC<br>Tyr        | AAT<br>Asn        | AAG<br>Lys        | · | 1344 |
| ATC<br>Ile        | GAA<br>Glu<br>450 | GGT<br>Gly        | CGT               | TGC<br>Cys        | GAT<br>Asp        | GGG<br>Gly<br>455 | GCA<br>Ala        | TTA<br>Leu        | AAT<br>Asn        | GAT<br>Asp        | TTA<br>Leu<br>460 | TGT<br>Cys        | ATC<br>Ile        | AAA<br>Lys        | GTT<br>Val        |   | 1392 |
| AAT<br>Asn<br>465 | AAT<br>Asn        | TGG<br>Trp        | GAC<br>Asp        | TTG<br>Leu        | TIT<br>Phe<br>470 | TTT<br>Phe        | AGT<br>Ser        | CCT<br>Pro        | TCA<br>Ser        | GAA<br>Glu<br>475 | GAT<br>Asp        | AAT<br>Asn        | TTT<br>Phe        | ACT<br>Thr        | AAT<br>Asn<br>480 |   | 1440 |
| gat<br>Asp        | CTA<br>Leu        | AAT<br>Asn        | AAA<br>Lys        | GGA<br>Gly<br>485 | GAA<br>Glu        | GAA<br>Glu        | ATT<br>Ile        | ACA<br>Thr        | TCT<br>Ser<br>490 | GAT<br>Asp        | ACT<br>Thr        | TAA<br>nea        | ATA<br>Ile        | GAA<br>Glu<br>495 | GCA.<br>Ala       |   | 1488 |
| GCA<br>Ala        | GAA<br>Glu        | GAA<br>Glu        | AAT<br>Asn<br>500 | ATT               | AGT<br>Ser        | TTA<br>Leu        | GAT<br>Asp        | TTA<br>Leu<br>505 | ATA<br>Ile        | CAA<br>Gln        | CAA<br>Gln        | TAT<br>Tyr        | TAT<br>Tyr<br>510 | TTA<br>Leu        | ACC<br>Thr        | · | 1536 |
| TTT               | AAT<br>Asn        | TTT<br>Phe<br>515 | GAT<br>Asp        | AAT<br>Asn        | GAA<br>Glu        | CCT<br>Pro        | GAA<br>Glu<br>520 | AAT<br>Asn        | ATT<br>Ile        | TCA<br>Ser        | ATA<br>Ile        | GAA<br>Glu<br>525 | AAT<br>Asn        | CTT<br>Leu        | TCA<br>Ser        |   | 1584 |
| AGT<br>Ser        | GAC<br>Asp<br>530 | ATT<br>Ile        | ATA<br>Ile        | GGC<br>Gly        | CAA<br>Gln        | TTA<br>Leu<br>535 | GAA<br>Glu        | CTT<br>Leu        | ATG<br>Met        | CCT<br>Pro        | AAT<br>Asn<br>540 | ATA<br>Ile        | GAA<br>Glu        | AGA<br>Arg        | TTT<br>Phe        | · | 1632 |
| CCT<br>Pro<br>545 | AAT<br>Asn        | GGA<br>Gly        | AAA<br>Lys        | AAG<br>Lys        | TAT<br>Tyr<br>550 | GAG<br>Glu        | TTA<br>Leu        | gat<br>Asp        | AAA<br>Lys        | TAT<br>Tyr<br>555 | ACT<br>Thr        | ATG<br>Met        | TTC<br>Phe        | CAT<br>His        | TAT<br>Tyr<br>560 |   | 1680 |
| CTT<br>Leu        | CGT               | GCT<br>Ala        | CAA<br>Gln        | GAA<br>Glu<br>565 | TTT<br>Phe        | GAA<br>Glu        | CAT<br>His        | GGT               | AAA<br>Lys<br>570 | TCT<br>Ser        | AGG<br>Arg        | ATT               | GCT<br>Ala        | TTA<br>Leu<br>575 | ACA<br>Thr        |   | 1728 |
| AAT<br>Asn        | TCT<br>Ser        | GTT<br>Val        | AAC<br>Asn<br>580 | GAA<br>Glu        | GCA<br>Ala        | TTA<br>Leu        | TTA<br>Leu        | AAT<br>Asn<br>585 | CCT<br>Pro        | AGT<br>Ser        | CGT<br>Arg        | GTT<br>Val        | TAT<br>Tyr<br>590 | ACA<br>Thr        | TTT<br>Phe        |   | 1776 |
| TTT<br>Phe        | TCT<br>Ser        | TCA<br>Ser<br>595 | GAC<br>Asp        | TAT<br>Tyr        | GTA<br>Val        | AAG<br>Lys        | AAA<br>Lys<br>600 | GTT<br>Val        | AAT<br>Asn        | AAA<br>Lys        | GCT<br>Ala        | ACG<br>Thr<br>605 | GAG<br>Glu        | GCA<br>Ala        | GCT<br>Ala        |   | 1824 |
| ATG<br>Met        | TTT<br>Phe<br>610 | TTA<br>Leu        | GGC<br>Gly        | TGG<br>Trp        | GTA<br>Val        | GAA<br>Glu<br>615 | CAA<br>Gln        | TTA<br>Leu        | GTA<br>Val        | TAT<br>Tyr        | GAT<br>Asp<br>620 | TTT<br>Phe        | ACC<br>Thr        | GAT<br>Asp        | GAA<br>Glu        |   | 1872 |
| ACT<br>Thr<br>625 | AGC<br>Ser        | GAA<br>Glu        | GTA<br>Val        | AGT<br>Ser        | ACT<br>Thr<br>630 | ACG<br>Thr        | GAT<br>Asp        | AAA<br>Lys        | ATT<br>Ile        | GCG<br>Ala<br>635 | GAT<br>Asp        | ATA<br>Ile        | ACT<br>Thr        | ATA<br>Ile        | ATT<br>Ile<br>640 |   | 1920 |

| AT<br>Il          | T CCI                 | A TAT             | T ATA             | GGA<br>Gly<br>645 | Pro               | GCT<br>Ala        | TT#               | A AAT<br>1 Asn    | TATA              | : Gly             | AA1<br>Asr        | ATC<br>Met        | TTI<br>Lei        | TA:<br>1 Ty:<br>659 | Lys               |   | 1968 |
|-------------------|-----------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---------------------|-------------------|---|------|
| GA<br>As          | T GAT<br>p Asp        | TTT<br>Phe        | GTA<br>Val<br>660 | GIA               | GCT<br>Ala        | TTA<br>Leu        | ATA<br>Ile        | TTT: Phe          | Ser               | GGA<br>Gly        | GCT<br>Ala        | GTI<br>Val        | AT1               | CTC                 | מידים             |   | 2016 |
| GA<br>G1          | A TTI<br>u Phe        | ATA               | Pro               | GAG<br>Glu        | ATT               | GCA<br>Ala        | ATA<br>Ile<br>680 | Pro               | GTA<br>Val        | TTA<br>Leu        | GGT<br>Gly        | ACT<br>Thr<br>685 | Phe               | GCA<br>Ala          | CTT<br>Leu        |   | 2064 |
| GT<br>Va          | A TCA<br>1 Ser<br>690 | Tyr               | ATT               | GCG<br>Ala        | AAT<br>Asn        | AAG<br>Lys<br>695 | GTT<br>Val        | CTA               | ACC<br>Thr        | GTT<br>Val        | CAA<br>Gln<br>700 | Thr               | ATA<br>Ile        | GAT<br>Asp          | AAT<br>Asn        |   | 2112 |
| GC<br>Al<br>70    | T TTA<br>a Leu<br>5   | AGT<br>Ser        | AAA<br>Lys        | AGA<br>Arg        | AAT<br>Asn<br>710 | GAA<br>Glu        | AAA<br>Lys        | TGG<br>Trp        | GAT<br>Asp        | GAG<br>Glu<br>715 | GTC<br>Val        | TAT<br>Tyr        | AAA<br>Lys        | TAT                 | ATA<br>Ile<br>720 |   | 2160 |
| GT.<br>Va         | A ACA<br>l Thr        | AAT<br>Asn        | TGG<br>Trp        | TTA<br>Leu<br>725 | GCA<br>Ala        | AAG<br>Lys        | GTT<br>Val        | AAT<br>Asn        | ACA<br>Thr<br>730 | CAG<br>Gln        | ATT               | GAT<br>Asp        | CTA<br>Leu        | ATA<br>Ile<br>735   | AGA<br>Arg        |   | 2208 |
| Lys               | A AAA<br>B Lys        | ATG<br>Met        | AAA<br>Lys<br>740 | GAA<br>Glu        | GCT<br>Ala        | TTA<br>Leu        | GAA<br>Glu        | AAT<br>Asn<br>745 | CAA<br>Gln        | GCA<br>Ala        | GAA<br>Glu        | GCA<br>Ala        | ACA<br>Thr<br>750 | AAG<br>Lys          | GCT<br>Ala        |   | 2256 |
| ATA               | A ATA                 | AAC<br>Asn<br>755 | TAT               | CAG<br>Gln        | TAT               | AAT<br>Asn        | CAA<br>Gln<br>760 | TAT               | ACT<br>Thr        | GAG<br>Glu        | GAA<br>Glu        | GAG<br>Glu<br>765 | AAA<br>Lys        | AAT<br>Asn          | AAT<br>Asn        | • | 2304 |
| ·Ile              | AAT<br>Asn<br>770     | Phe               | Asn               | Ile               | Asp               | 775               | Leu               | Ser               | Ser               | Lys               | Leu<br>780        | Asn               | Glu               | Ser                 | Ile               |   | 2352 |
| 785               |                       | Ala               | Met               | Ile               | 790               | Ile               | Asn               | Lys               | Phe               | Leu<br>795        | Asn               | Gln               | Cys               | Ser                 | Val<br>800        | • | 2400 |
| Ser               | TAT                   | Leu               | Met               | 805               | Ser               | Met               | Ile               | Pro               | Tyr<br>810        | Gly               | Val               | Lys               | Arg               | Leu<br>815          | Glu               |   | 2448 |
| GA1<br>Asp        | Phe                   | GAT<br>Asp        | GCT<br>Ala<br>820 | AGT<br>Ser        | CTT<br>Leu        | AAA<br>Lys        | Asp               | GCA<br>Ala<br>825 | TTA<br>Leu        | TTA<br>Leu        | AAG<br>Lys        | TAT<br>Tyr        | ATA<br>Ile<br>830 | TAT<br>Tyr          | GAT<br>Asp        |   | 2496 |
| Asn               | AGA<br>Arg            | Gly<br>835        | Thr               | Leu               | Ile               | Gly               | Gln<br>840        | Val               | Asp               | Arg               | Leu               | Lys<br>845        | Asp               | Lys                 | Val               |   | 2544 |
| Asn               | AAT<br>Asn<br>850     | Thr               | Leu               | Ser               | Thr               | <b>Asp</b><br>855 | Ile               | Pro               | Phe               | Gln               | Leu<br>860        | Ser               | Lys               | Tyr                 | Val               |   | 2592 |
| GAT<br>Asp<br>865 | AAT                   | CAA<br>Gln        | AGA<br>Arg        | TTA<br>Leu        | TTA<br>Leu<br>870 | TCT<br>Ser        | ACA<br>Thr        | TTT<br>Phe        | ACT .<br>Thr      | GAA<br>Glu<br>875 | TAT<br>Tyr        | ATT<br>Ile        | AAG<br>Lys        | TCT<br>Ser          | AGG<br>Arg<br>880 |   | 2640 |
| CCI               | CAA<br>Gln            | TCT<br>Ser        | Lys               | GTT<br>Val<br>885 | AAA<br>Lys        | AGA<br>Arg        | CAA<br>Gln        | ATA<br>Ile        | TTT<br>Phe<br>890 | TCA<br>Ser        | GGC<br>Gly        | TAT<br>Tyr        | CAA<br>Gln        | TCT<br>Ser<br>895   | GAT<br>Asp        |   | 2688 |
|                   | GAT<br>Asp            |                   |                   |                   |                   |                   |                   |                   |                   |                   | TGA<br>*          |                   |                   |                     |                   |   | 2724 |

#### (2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 908 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly
1 5 10

Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro

Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg

Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu 50 60

Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr
65 70 75 80

. Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu 85 90 95

Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val

Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys

Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr 130 135 140

Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile 145 150 155 160

Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr 165 170 175

Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe 180 185 190

Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu 195 200 205

Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu 210 215 220

Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn 225 230 240

Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu 245 250 255

Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys 260 265 270

Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn 275 280 285

Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val 290 295 300



- 78 -

Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu 325 Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn 395 Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu 410 Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys Ile Glu Gly Arg Cys Asp Gly Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser 525 Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr 550 Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu 615 Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys

WO 98/07864 PCT/GB97/02273

- 79 -

Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu 660 665

Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu 675 680 685

Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn 690 695 700

Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile
705 710 715 720

Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg
725 730 735

Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala
740 745 750

Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn 755

Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile
770 780

Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val 785 790 795 800

Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu 805 810 815

Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp 820 825 830

Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val

Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val

Asp Asn Gln Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys Ser Arg 865 870 870 880

Pro Gln Ser Lys Val Lys Arg Gln Ile Phe Ser Gly Tyr Gln Ser Asp 885 890 895

Ile Asp Thr His Asn Arg Ile Lys Asp Glu Leu 900 905

#### (2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3042 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION:1..3042
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATG CAG TTC GTG AAC AAG CAG TTC AAC TAT AAG GAC CCT GTA AAC GGT Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly 1

48

|          |            |                   | 2              | 0           |            | e пу:             |                       | 2       | o Asi<br>5   | n Al           | a Gl       | y 'Gli                | n Me<br>3      | t Gl<br>O  | G CCC                | •   | 96  |
|----------|------------|-------------------|----------------|-------------|------------|-------------------|-----------------------|---------|--------------|----------------|------------|-----------------------|----------------|------------|----------------------|-----|-----|
|          | ,          | 3                 | 5              | c a,        | 5 11       | e ure             | 4 (                   | о гу:   | 2 TT6        | ? Trj          | p Va       | 1 Ile<br>45           | e Pro<br>5     | o Gl       | A CGC<br>u Arg       |     | 144 |
|          | 5          | 0                 |                | r no        |            | 55                | GIL                   | ı GT    | / Asr        | Let            | 1 As:      | n Pro<br>O            | Pro            | o Pr       | G GAA<br>o Glu       |     | 192 |
| 65       | 5          |                   |                |             | 70         | )                 | ıyı                   | . lyr   | Asp          | 75             | Th         | r Tyr                 | : Leu          | ı Se       | C ACA<br>r Thr<br>80 |     | 240 |
|          | <i>'</i>   |                   | <b>- 1</b> 37. | 85          | y war      | ı ışr             | Leu                   | туу     | 90<br>G1A    | Val            | Thi        | Lys                   | Leu            | Phe<br>95  | -                    |     | 288 |
| <b>3</b> |            |                   | 100            | )           |            | , nea             | GIY                   | 105     | met          | Leu            | Leu        | Thr                   | Ser<br>110     | Ile        | GTC<br>Val           | ÷   | 336 |
| 3        |            | 115               | 5              | , , , , , , | ilp        | GIY               | 120                   | ser     | Thr          | Ile            | Asp        | Thr<br>125            | Glu            | Leu        | AAG<br>Lys           | •   | 384 |
|          | 130        |                   |                | ASII        | Cys        | ATT<br>Ile<br>135 | ASN                   | val     | iie          | Gln            | Pro<br>140 | Asp                   | Gly            | Ser        | Tyr                  |     | 432 |
| 145      |            | 020               |                | neu         | 150        | CTC<br>Leu        | AdT                   | TTE     | ite          | G1y<br>155     | Pro        | Ser                   | Ala            | Asp        | Ile<br>160           |     | 480 |
|          | <b>511</b> | 2 110             | GIU            | 165         | пув        | AGC<br>Ser        | Pne                   | GIÀ     | H15<br>170   | Glu            | Val        | Leu                   | Asn            | Leu<br>175 | Thr                  |     | 528 |
|          |            | ,                 | 180            | Cly         | Jei        |                   | GIN                   | 185     | i            | Arg            | Phe        | Ser                   | Pro<br>190     | Asp        | Phe                  |     | 576 |
|          | - 110      | 195               | FIIC           | Gra         | GIU        |                   | 200                   | GIU     | Val          | Asp            | Thr        | Asn<br>205            | Pro            | Leu-       | Leu                  | . • | 624 |
| 0_1      | 210        | Cly               | nys            | FIIG        | WIG        | ACT<br>Thr<br>215 | Asp                   | Pro     | Ala          | Val            | Thr<br>220 | Leu                   | Ala            | His        | Glu .                |     | 672 |
| 225      |            |                   | . ATG          | GIY         | 230        | CGT<br>Arg        | Leu                   | lyr     | GIÀ          | 11e<br>235     | Ala        | Ile                   | Asn            | Pro        | Asn<br>240           |     | 720 |
|          | <b>741</b> | F11C              | Lys            | 245         | ABII       | ACC I             | ASD .                 | YTA .   | Tyr '<br>250 | Tyr            | Glu        | Met :                 | Ser            | Gly<br>255 | Leu                  |     | 768 |
| ·        |            | SET               | 260            | GIU         |            | CTG (<br>Leu )    | Arg                   | 7nr 265 | Phe (        | Gly (          | Gly        | His A                 | Asp .<br>270   | Ala        | Lys                  |     | 816 |
| Phe      |            | GAC<br>Asp<br>275 | AGC<br>Ser     | TTG<br>Leu  | CAG<br>Gln | GAG A             | AAC (<br>Asn (<br>280 | GAG 1   | Phe I        | CGT (<br>Arg ) | Leu        | TAC 1<br>Tyr 1<br>285 | TAC '<br>Tyr ' | TAC<br>Tyr | AAC<br>Asn           |     | 864 |

- 81 -

|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | · • ·             |                   |                   |                   |                   |                   |      |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|
| AAG<br>Lys        | TTT<br>Phe<br>290 | AAA<br>Lys        | GAT<br>Asp        | ATT<br>Ile        | GCA<br>Ala        | AGT<br>Ser<br>295 | ACA<br>Thr        | CTG<br>Leu        | AAC<br>Asn        | AAG<br>Lys        | GCT<br>Ala<br>300 | AAG<br>Lys        | TCC<br>Ser        | ATT<br>Ile        | GTG<br>Val        | 912  |
| GGT<br>Gly<br>305 | ACC<br>Thr        | ACT<br>Thr        | GCT<br>Ala        | TCA<br>Ser        | TTA<br>Leu<br>310 | CAG<br>Gln        | TAT               | ATG<br>Met        | AAA<br>Lys        | AAT<br>Asn<br>315 | GTT<br>Val        | TTT<br>Phe        | AAA<br>Lys        | GAG<br>Glu        | AAA<br>Lys<br>320 | 960  |
| TAT<br>Tyr        | CTC<br>Leu        | CTA<br>Leu        | TCT<br>Ser        | GAA<br>Glu<br>325 | GAT<br>Asp        | ACA<br>Thr        | TCT<br>Ser        | GGA<br>Gly        | AAA<br>Lys<br>330 | TTT<br>Phe        | TCG<br>Ser        | GTA<br>Val        | GAT<br>Asp        | AAA<br>Lys<br>335 | TTA<br>Leu        | 1008 |
| AAA<br>Lys        | TTT<br>Phe        | GAT<br>Asp        | AAG<br>Lys<br>340 | TTA<br>Leu        | TAC<br>Tyr        | AAA<br>Lys        | ATG<br>Met        | TTA<br>Leu<br>345 | ACA<br>Thr        | GAG<br>Glu        | ATT<br>Ile        | TAC<br>Tyr        | ACA<br>Thr<br>350 | GAG<br>Glu        | GAT<br>Asp        | 1056 |
| AAT<br>Asn        | TTT<br>Phe        | GTT<br>Val<br>355 | AAG<br>Lys        | TTT<br>Phe        | TTT<br>Phe        | AAA<br>Lys        | GTA<br>Val<br>360 | CTT               | AAC<br>Asn        | AGA<br>Arg        | AAA<br>Lys        | ACA<br>Thr<br>365 | TAT<br>Tyr        | TTG<br>Leu        | AAT<br>Asn        | 1104 |
| TTT<br>Phe        | GAT<br>Asp<br>370 | AAA<br>Lys        | GCC<br>Ala        | GTA<br>Val        | TTT               | AAG<br>Lys<br>375 | ATA<br>Ile        | AAT<br>Asn        | ATA<br>Ile        | GTA<br>Val        | CCT<br>Pro<br>380 | <u>AAG</u><br>Lys | <u>GTA</u><br>Val | AAT<br>Asn        | TAC<br>Tyr        | 1152 |
| ACA<br>Thr<br>385 | ATA<br>Ile        | TAT<br>Tyr        | GAT<br>Asp        | GGA<br>Gly        | TTT<br>Phe<br>390 | AAT<br>Asn        | TTA<br>Leu        | AGA<br>Arg        | AAT<br>Asn        | ACA<br>Thr<br>395 | TAA<br>neA        | TTA<br>Leu        | GCA<br>Ala        | GCA<br>Ala        | AAC<br>Asn<br>400 | 1200 |
| TTT<br>Phe        | AAT<br>Asn        | GGT<br>Gly        | CAA<br>Gln        | AAT<br>Asn<br>405 | ACA<br>Thr        | GAA<br>Glu        | ATT<br>Ile        | AAT<br>Asn        | AAT<br>Asn<br>410 | ATG<br>Met        | AAT<br>Asn        | TTT               | ACT<br>Thr        | AAA<br>Lys<br>415 | CTA.<br>Leu       | 1248 |
| AAA<br>Lys        | AAT<br>Asn        | TTT<br>Phe        | ACT<br>Thr<br>420 | GGA<br>Gly        | TTG<br>Leu        | TTT<br>Phe        | GAA<br>Glu        | TTT<br>Phe<br>425 | TAT<br>Tyr        | AAG<br>Lys        | TTG<br>Leu        | CTA<br>Leu        | TGT<br>Cys<br>430 | GTA<br>Val        | AGA<br>Arg        | 1296 |
| GGG<br>Gly        | ATA<br>Ile        | ATA<br>Ile<br>435 | ACT<br>Thr        | TCT<br>Ser        | AAA<br>Lys        | ACT<br>Thr        | AAA<br>Lys<br>440 | TCA<br>Ser        | TTA<br>Leu        | GAT<br>Asp        | AAA<br>Lys        | GGA<br>Gly<br>445 | TAC<br>Tyr        | AAT<br>Asn        | AAG<br>Lys        | 1344 |
| ATC<br>Ile        | GAA<br>Glu<br>450 | GGT<br>Gly        | CGT<br>Arg        | TGC<br>Cys        | GAT<br>Asp        | GGG<br>Gly<br>455 | GCA<br>Ala        | TTA<br>Leu        | AAT<br>Asn        | GAT<br>Asp        | TTA<br>Leu<br>460 | TGT<br>Cys        | ATC<br>Ile        | AAA<br>Lys        | GTT<br>Val        | 1392 |
| AAT<br>Asn<br>465 | AAT<br>Asn        | TGG<br>Trp        | GAC<br>Asp        | TTG<br>Leu        | TTT<br>Phe<br>470 | TTT               | AGT<br>Ser        | CCT<br>Pro        | TCA<br>Ser        | GAA<br>Glu<br>475 | GAT<br>Asp        | AAT<br>Asn        | TTT               | ACT<br>Thr        | AAT<br>Asn<br>480 | 1440 |
| GAT<br>Asp        | CTA<br>Leu        | AAT<br>Asn        | Lys<br>Lys        | GGA<br>Gly<br>485 | GAA.<br>Glu       | GAA<br>Glu        | ATT Ile           | ACA<br>Thr        | TCT<br>Ser<br>490 | GAT<br>Asp        | ACT<br>Thr        | AAT<br>Asn        | ATA<br>Ile        | GAA<br>Glu<br>495 | GCA<br>Ala        | 1488 |
| GCA<br>Ala        | GAA<br>Glu        | GAA<br>Glu        | AAT<br>Asn<br>500 | ATT<br>Ile        | AGT<br>Ser        | TTA<br>Leu        | GAT<br>Asp        | TTA<br>Leu<br>505 | ATA<br>Ile        | CAA<br>Gln        | CAA<br>Gln        | TAT<br>Tyr        | TAT<br>Tyr<br>510 | TTA<br>Leu        | ACC               | 1536 |
| TTT               | AAT<br>Asn        | TTT<br>Phe<br>515 | GAT<br>Asp        | AAT<br>Asn        | GAA<br>Glu        | CCT<br>Pro        | GAA<br>Glu<br>520 | AAT<br>Asn        | ATT<br>Ile        | TCA<br>Ser        | ATA<br>Ile        | GAA<br>Glu<br>525 | AAT<br>Asn        | CTT               | TCA<br>Ser        | 1584 |
| Ser               | Asp<br>530        | Ile               | Ile               | Gly               | Gln               | TTA<br>Leu<br>535 | Glu               | Leu               | Met               | Pro               | Asn<br>540        | Ile               | Glu               | Arg               | Phe               | 1632 |
| CCT<br>Pro<br>545 | TAA<br>neA        | GGA<br>Gly        | AAA<br>Lys        | AAG<br>Lys        | TAT<br>Tyr<br>550 | GAG<br>Glu        | TTA<br>Leu        | GAT<br>Asp        | AAA<br>Lys        | TAT<br>Tyr<br>555 | ACT               | ATG<br>Met        | TTC<br>Phe        | CAT<br>His        | TAT<br>Tyr<br>560 | 1680 |



- 82.-

|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | •                 |      |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|
| CTT<br>Leu        | CGT               | GCT<br>Ala        | CAA<br>Gln        | GAA<br>Glu<br>565 | TTT               | GAA<br>Glu        | CAT               | GGT<br>Gly        | AAA<br>Lys<br>570 | TCT               | AGG<br>Arg        | ATT<br>Ile        | GCT<br>Ala        | TTA<br>Leu<br>575 | ACA,<br>Thr       | 1728 |
| AAT<br>Asn        | TCT<br>Ser        | GTT<br>Val        | AAC<br>Asn<br>580 | GAA<br>Glu        | GCA<br>Ala        | TTA<br>Leu        | Leu               | AAT<br>Asn<br>585 | CCT<br>Pro        | AGT<br>Ser        | CGT<br>Arg        | GTT<br>Val        | TAT<br>Tyr<br>590 | ACA<br>Thr        | TTT<br>Phe        | 1776 |
| TTT<br>Phe        | TCT               | TCA<br>Ser<br>595 | GAC<br>Asp        | TAT               | GTA<br>Val        | AAG<br>Lys        | AAA<br>Lys<br>600 | GTT<br>Val        | AAT<br>Asn        | AAA<br>Lys        | GCT<br>Ala        | ACG<br>Thr<br>605 | GAG<br>Glu        | GCA<br>Ala        | GCT<br>Ala        | 1824 |
| ATG<br>Met        | TTT<br>Phe<br>610 | TTA<br>Leu        | GGC<br>Gly        | TGG<br>Trp        | GTA<br>Val        | GAA<br>Glu<br>615 | CAA<br>Gln        | TTA<br>Leu        | GTA<br>Val        | TAT               | GAT<br>Asp<br>620 | TTT<br>Phe        | ACC<br>Thr        | GAT<br>Asp        | GAA<br>Glu        | 1872 |
| ACT<br>Thr<br>625 | AGC<br>Ser        | GAA<br>Glu        | GTA<br>Val        | AGT<br>Ser        | ACT<br>Thr<br>630 | ACG<br>Thr        | GAT<br>Asp        | AAA<br>Lys        | ATT<br>Ile        | GCG<br>Ala<br>635 | GAT<br>Asp        | ATA<br>Ile        | ACT               | ATA<br>Ile        | ATT<br>Ile<br>640 | 1920 |
| ATT               | CCA<br>Pro        | TAT<br>Tyr        | ATA<br>Ile        | GGA<br>Gly<br>645 | CCT               | GCT<br>Ala        | TTA<br>Leu        | AAT<br>Asn        | ATA<br>Ile<br>650 | GGT.<br>Gly       | AAT<br>Asn        | ATG<br>Met        | TTA<br>Leu        | TAT<br>Tyr<br>655 | AAA<br>Lys        | 1968 |
| GAT<br>Asp        | GAT<br>Asp        | TTT<br>Phe        | GTA<br>Val<br>660 | GGT<br>Gly        | GCT<br>Ala        | TTA<br>Leu        | ATA<br>Ile        | TTT<br>Phe<br>665 | TCA<br>Ser        | GGA<br>Gly        | GCT<br>Ala        | GTT<br>Val        | ATT<br>Ile<br>670 | CTG<br>Leu        | TTA<br>Leu        | 2016 |
| GAA<br>Glu        | TTT<br>Phe        | ATA<br>Ile<br>675 | CCA<br>Pro        | GAG<br>Glu        | ATT               | GCA<br>Ala        | ATA<br>Ile<br>680 | CCT<br>Pro        | GTA<br>Val        | TTA<br>Leu        | GGT<br>Gly        | ACT<br>Thr<br>685 | TTT<br>Phe        | GCA<br>Ala        | CTT<br>Leu        | 2064 |
| Val               | TCA<br>Ser<br>690 | TAT<br>Tyr        | ATT Ile           | GCG<br>Ala        | TAA<br>Asn        | AAG<br>Lys<br>695 | GTT<br>Val        | CTA<br>Leu        | ACC               | GTT<br>Val        | CAA<br>Gln<br>700 | ACA<br>Thr        | ATA<br>Ile        | GAT<br>Asp        | AAT<br>Asn        | 2112 |
| GCT<br>Ala<br>705 | TTA<br>Leu        | AGT<br>Ser        | AAA<br>Lys        | AGA<br>Arg        | AAT<br>Asn<br>710 | GAA<br>Glu        | AAA<br>Lys        | TGG<br>Trp        | GAT<br>Asp        | GAG<br>Glu<br>715 | GTC<br>Val        | TAT<br>Tyr        | AAA<br>Lys        | TAT<br>Tyr        | ATA<br>Ile<br>720 | 2160 |
| GTA<br>Val        | ACÀ<br>Thr        | AAT<br>Asn        | TGG<br>Trp        | TTA<br>Leu<br>725 | GCA<br>Ala        | AAG<br>Lys        | GTT<br>Val        | AAT<br>Asn        | ACA<br>Thr<br>730 | CAG<br>Gln        | ATT<br>Ile        | GAT<br>Asp        | CTA<br>Leu        | ATA<br>Ile<br>735 | AGA<br>Arg        | 2208 |
| AAA<br>Lys        | AAA<br>Lys        | ATG<br>Met        | AAA<br>Lys<br>740 | GAA<br>Glu        | GCT<br>Ala        | TTA<br>Leu        | GAA<br>Glu        | AAT<br>Asn<br>745 | CAA<br>Gln        | GCA<br>Ala        | GAA<br>Glu        | GCA<br>Ala        | ACA<br>Thr<br>750 | AAG<br>Lys        | GCT<br>Ala        | 2256 |
| ATA<br>Ile        | ATA<br>Ile        | AAC<br>Asn<br>755 | TAT<br>Tyr        | CAG<br>Gln        | TAT<br>Tyr        | AAT<br>Asn        | CAA<br>Gln<br>760 | TAT<br>Tyr        | ACT<br>Thr        | GAG<br>Glu        | GAA<br>Glu        | GAG<br>Glu<br>765 | AAA<br>Lys        | AAT<br>Asn        | AAT<br>Asn        | 2304 |
| Ile               | AAT<br>Asn<br>770 | TTT<br>Phe        | AAT<br>Asn        | ATT<br>Ile        | GAT<br>Asp        | GAT<br>Asp<br>775 | TTA<br>Leu        | AGT<br>Ser        | TCG<br>Ser        | AAA<br>Lys        | CTT<br>Leu<br>780 | AAT<br>Asn        | GAG<br>Glu        | TCT<br>Ser        | ATA<br>Ile        | 2352 |
|                   |                   |                   | ATG<br>Met        |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | 2400 |
|                   |                   |                   | ATG<br>Met        |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | 2448 |
|                   |                   |                   | GCT<br>Ala<br>820 |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | 2496 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |      |

- 83 -

|              |                     |                     |            |                |            |                   |                    |            |                |            |                   |                       |            |            |                    | •    |
|--------------|---------------------|---------------------|------------|----------------|------------|-------------------|--------------------|------------|----------------|------------|-------------------|-----------------------|------------|------------|--------------------|------|
| AA'<br>As:   | T AGA<br>n Arg      | A GG<br>g Gly<br>83 | A TIN      | r TT/<br>r Lei | A AT       | r GG7<br>e Gly    | CAA<br>Glr<br>840  | ı val      | A GAT<br>L Asp | r AGA      | A TT              | A AAA<br>u Lys<br>845 | : Asp      | Lys        | GTT<br>Val         | 2544 |
| AA'<br>Ası   | TAA T<br>RAS<br>RSS | r Tiis              | A CTT      | r AG7<br>1 Sez | T AC       | GAT<br>Asp<br>855 | ile                | CCI<br>Pro | TTT<br>Phe     | CAG<br>Glr | CT:<br>Lei<br>860 | ı Ser                 | Lys        | TAC<br>Tyr | GTA<br>Val         | 2592 |
| 869          | o Asi               | i Gir               | Arg        | Leu            | 870        | ser               | Thr                | · Phe      | Thr            | 875        | Туз               | ATT                   | Lys        | Ser        | Gly<br>880         | 2640 |
| Let          | . ASD               | ser                 | PIO        | 885            | ATS        | AIA               | HIS                | Tyr        | 890            | Gln        | His               | GAT<br>Asp            | Glu        | Ala<br>895 | Val                | 2688 |
| wah          | ASII                | гуз                 | 900        | ASI            | гÀ2        | GIU               | Gin                | 905        | Asn            | Ala        | Phe               | TAT                   | Glu<br>910 | Ile        | Leụ                | 2736 |
| nrs          | Deu                 | 915                 | ASI        | reu            | Asn        | GIU               | 920                | Gin        | Arg            | Asn        | Ala               | TTC<br>Phe<br>925     | Ile        | Gln        | Ser                | 2784 |
| Dea          | 930                 | Asp                 | Asp        | Pro            | ser        | 935               | Ser                | Ala        | Asn            | Leu        | Leu<br>940        | GCA<br>Ala            | Glu        | Ala        | Lys                | 2832 |
| 9 <b>4</b> 5 | Leu                 | ASI                 | Asp        | Ala            | 950        | ATA               | Pro                | Lys        | Val            | Asp<br>955 | Asn               | AAA<br>Lys            | Phe        | Asn        | Lys<br>9 <b>60</b> | 2880 |
| GIU          | GIII                | GIN                 | ASI        | 965            | Pue        | iyr               | Glu                | Ile        | 1eu<br>970     | His        | Leu               | CCT<br>Pro            | Asn        | Leu<br>975 | Asn                | 2928 |
| GIU          | GIU                 | GIN                 | 980        | ASN            | AIA        | Pne               | IIe                | 985        | Ser            | Leu        | Lys               |                       | Asp<br>990 | Pro        | Ser                | 2976 |
| CAA<br>Gln   | AGC<br>Ser          | GCT<br>Ala<br>995   | AAC<br>Asn | CTT<br>Leu     | TTA<br>Leu | ALA (             | GAA<br>Glu<br>1000 | Ala        | AAA<br>Lys     | AAG<br>Lys | CTA<br>Leu        | AAT<br>Asn<br>1005    | GAT<br>Asp | GCT<br>Ala | CAG<br>Gln         | 3024 |
| Ala          | CCG<br>Pro<br>1010  | Lys                 |            |                | TAG        |                   | •                  |            |                |            |                   |                       |            |            |                    | 3042 |

#### (2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1014 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly 10

Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro

Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg



. 84 -

| Asp        | Thr<br>50  | Phe        | Thr        | Asn         | Pro        | Glu<br>55  | Glu        | Gly        | Asp        | Leu                | Asn<br>60  | Pro        | Pro        | Pro        | Glu        |
|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|--------------------|------------|------------|------------|------------|------------|
| Ala<br>65  | Lys        | Gln        | Val        | Pro         | Val<br>70  | Ser        | Tyr        | Tyr        | Asp        | Ser<br>75          | Thr        | Tyr        | Leu        | Ser        | Thr<br>80  |
| Asp        | Asn        | Glu        | Lys        | Asp<br>85   | Asn        | Tyr        | Leu        | Lys        | Gly<br>90  | Val                | Thr        | Lys        | Leu        | Phe<br>95  | Glu        |
| Arg        | Ile        | Tyr        | Ser<br>100 | Thr         | Asp        | Leu        | Gly        | Arg<br>105 | Met        | Leu                | Leu        | Thr        | Ser<br>110 | Ile        | Val        |
| Arg        | Gly        | Ile<br>115 | Pro        | Phe         | Trp        | Gly        | Gly<br>120 |            | Thr        | Ile                | Asp        | Thr<br>125 | Glu        | Leu        | Lys        |
| Val        | Ile<br>130 | Asp        | Thr        | Asn         | Cys        | Ile<br>135 | Asn        | Val        | Ile        | Gln                | Pro<br>140 | Asp        | Gly        | Ser        | Tyr        |
| Arg<br>145 | Ser        | Glu        | Glu        | Leu         | Asn<br>150 | Leu        | Val        | Ile        | Ile        | Gly<br>155         | Pro        | Ser        | Ala        | Asp        | Ile<br>160 |
| 'lle       | Gln        | Phe        | Glu        | Cys<br>165  | Lys        | Ser        | Phe        | Gly        | His<br>170 | Glu                | Val        | Leu        | Asn        | Leu<br>175 | Thr        |
| Arg        | Asn        | Gly        | Tyr<br>180 | Gly         | Ser        | Thr        | Gln        | Tyr<br>185 | Ile        | Arg                | Phe        | Ser        | Pro<br>190 | Asp        | Phe        |
| Thr        | Phe        | Gly<br>195 | Phe        | Glu         | Glu        | Ser        | Leu<br>200 | Glu        | Val        | Asp                | Thr        | Asn<br>205 | Pro        | Leu        | Leu        |
| Gly        | Ala<br>210 | Gly        | Lys        | Phe         | Ala        | Thr<br>215 | Asp        | Pro        | Ala        | Val                | Thr 220    | Leu        | Ala        | His        | Glu        |
| Leu<br>225 | Ile        | His        | Ala        | Gly         | His<br>230 | Arg        | Leu        | Tyr        | Gly        | Ile<br>235         | Ala        | Ile        | Asn        | Pro        | Asn<br>240 |
| Arg        | Val        | Phe        | Lys        | Val<br>245  | Asn        | Thr        | Asn        | Ala        | Tyr<br>250 | Tyr                | Glu        | Met        | Ser        | Gly<br>255 | Leu        |
| Glu        | Val        | Ser        | Phe<br>260 | Glu         | Glu        | Leu        | Arg        | Thr<br>265 | Phe        | Gly                | Gly.       | His        | Asp<br>270 | Ala        | Lys        |
| Phe        | Ile        | Asp<br>275 | Ser        | Leu.        | Gln        | Glu        | Asn<br>280 | Glu        | Phe        | Arg                | Leu        | Tyr<br>285 | Tyr        | Tyr        | Asn        |
| Lys        | Phe 290    | Lys        | Asp        | Ile         | Ala        | Ser<br>295 | Thr        | Leu        | Asn        | Lys                | Ala<br>300 | Lys        | Ser        | Ile        | Val        |
| Gly<br>305 | Thr        | Thr        | Ala        | Ser         | Leu<br>310 | Gln        | Tyr        | Met        | Lys        | Дв <u>й</u><br>315 | Val        | Phe        | Lys        | Glu        | Lys<br>320 |
| Tyr        | Leu        | Leu        | Ser        | Glu<br>.325 | Asp        | Thr        | Ser        | Gly        | Lys        | Phe                | Ser        | Val        | Asp        | Lys<br>335 | Leu        |
| Lys        | Phe        | Asp        | Lys<br>340 | Leu         | Tyr        | Lys        | Met        | Leu<br>345 | Thr        | Glu                | Ile        | Tyr        | Thr<br>350 | Glu        | Asp        |
| Asn        | Phe        | Val<br>355 | Lys        | Phe         | Phe        | Lys        | Val<br>360 | Leu        | Asn        | Arg                | Lys        | Thr<br>365 | Tyr        | Leu        | Asn        |
| Phe        | Asp<br>370 | Lys        | Ala        | Val         | Phe        | Lys<br>375 | Ile        | Asn        | Ile        | Val                | Pro<br>380 | Lys        | Val        | Asn        | Tyr        |
| Thr<br>385 | Ile        | Tyr        | Asp        | Gly         | Phe 390    | Asn        | Leu        | Arg        | Asn        | Thr<br>395         | Asn        | Leu        | Ala        | Ala        | Asn<br>400 |

- 85 -

Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu 410 Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg 425 Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys Ile Glu Gly Arg Cys Asp Gly Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr 550 Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr 565 Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile 630 Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu 680 Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala



- 86 -

Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val 855 Asp Asn Gln Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys Ser Gly Leu Asn Ser Pro Gly Ala Ala His Tyr Ala Gln His Asp Glu Ala Val Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile Leu 905 His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser 915 Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys 935 Lys Leu Asn Asp Ala Gln Ala Pro Lys Val Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile Leu His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys Lys Leu Asn Asp Ala Gln 1000 995 Ala Pro Lys Val Asp \* 1010

- (2) INFORMATION FOR SEQ ID NO: 19:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 3509 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: double (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (ix) FEATURE:
    - (A) NAME/KEY: CDS
    - (B) LOCATION:1..3509
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

- 87 -

|   | •                 |                   |              |                  |                      |                    |                   |                   |                      |                    |                |                  | _                | •                 |                   |                    |                  |                   |            |                  |         |
|---|-------------------|-------------------|--------------|------------------|----------------------|--------------------|-------------------|-------------------|----------------------|--------------------|----------------|------------------|------------------|-------------------|-------------------|--------------------|------------------|-------------------|------------|------------------|---------|
|   | 7-10              | 1                 | 10           | VAI              | 1111                 | . 11               | 5 A               | an A              | AT T<br>sn P         | ne A               | sn             | Ty:              | As:              | n As              | p P               | ro                 | Ile              | As                | sp<br>L5   | Asn              | 48      |
|   | AA<br>As          | T A               | AT<br>sn     | ATT<br>Ile       | ATT<br>Ile<br>20     | : ME               | G AT              | G G               | AG C                 | CT C               | CA<br>ro<br>25 | TTT<br>Phe       | GCC<br>Ala       | G AG              | ia G              | GT A               | ACG<br>Thr<br>30 | G]                | ig<br>Ly   | AGA<br>Arg       | 96      |
|   | TA<br>Ty          | T T               | AT .<br>yr : | AAA<br>Lys<br>35 | GCT                  | Ph                 | T AA<br>e Ly      | A A'<br>'s I      | IC A                 | CA G<br>hr A<br>10 | AT<br>sp       | CGI<br>Arg       | AT?              | TG<br>Tr          | p I               | ra A<br>le 1<br>15 | ATA<br>(le       | Pr                | G<br>O     | GAA<br>Glu       | 144     |
|   | AG.               | 9 1               | AT I         | ACT<br>Thr       | TTT<br>Phe           | GG)                | A TA<br>/ Ty      | ı rı              | VA CO                | CT G.              | AG<br>lu       | GAT<br>Asp       | TTT              | AA :<br>As:       | n Ly              | VA A               | GT<br>Ser        | TC<br>Se          | c (        | GGT<br>Gly       | 192     |
|   | AT<br>Ile<br>6    | e Fi              | rr /<br>ie / | AAT<br>Asn       | AGA<br>Arg           | GA1<br>Asp         | GT<br>Va<br>7     | T C               | T GA<br>'s Gl        | LA TI              | AT<br>/r       | TAT<br>Tyr       | GAT<br>Asp<br>75 | Pro               | A GA<br>O As      | T T                | AC<br>Yr         | TT                | A A        | AAT<br>Asn<br>80 | 240     |
| 1 | ACT<br>Thi        | r Aa<br>c As      | T C          | SAT              | FAY<br>TÀS           | AAG<br>Lys<br>85   | ASI               | r AT              | A TI<br>e Ph         | T TI<br>e Le       | ra (           | CAA<br>Gln<br>90 | ACA<br>Thr       | ATO               | G AT              | C A<br>e L         | AG<br>YS         | TT/<br>Lev<br>95  | ı F        | TT               | 288     |
|   | AA1<br>Asr        | AG<br>Ar          | A A<br>g I   | TC               | AAA<br>Lys<br>100    | TCA<br>Ser         | AAJ<br>Lys        | A CC              | A TT<br>o Le         | G GG<br>u Gl<br>10 | y (            | GAA<br>Glu       | AAG<br>Lys       | TTA               | TT.               | u G                | AG<br>lu<br>10   | ATC<br>Met        | G A        | TT<br>le         | 336     |
|   | ATA<br>Ile        | AA A<br>RA :      | 11 6         | GT<br>1y<br>15   | ATA<br>Ile           | CCT<br>Pro         | TAT               | CT<br>Le          | T GG.<br>u Gl:<br>12 | y As               | T P            | AGA<br>Arg       | CGT<br>Arg       | GTT<br>Val        | CC<br>Pro         | o Le               | rc<br>eu         | GAA<br>Glu        | G<br>G     | AG<br>lu         | <br>384 |
|   | TTT               | AA<br>As:         |              | CA I             | AAC<br>Asn           | ATT<br>Ile         | GCT<br>Ala        | AG:<br>Se:<br>139 | r GT<br>r Val        | A AC<br>l Th       | T G            | TT<br>Val        | AAT<br>Asn       | AAA<br>Lys<br>140 | Let               | A AT               | rc .<br>le :     | AGT<br>Ser        | A          | AT<br>sn         | 432     |
|   | CCA<br>Pro<br>145 | GI                | A G          | Lu '             | GTG<br>Val           | GIU                | Arg               | Lys               | A AAA<br>E Lys       | GG<br>GI           | y I            | le               | Phe              | GCA<br>Ala        | AAT<br>Asr        | TI<br>Le           | ra z             | ATA<br>Ile        | A:<br>1:   | le               | 480     |
|   | TTT<br>Phe        | GG/<br>Gly        | A CO         | CT (             | TA.                  | CCA<br>Pro<br>165  | GTT<br>Val        | TTA<br>Lev        | AAT<br>Asn           | GAJ<br>Gli         | ı A            | AT (<br>sn (     | GAG<br>Glu       | ACT<br>Thr        | ATA               | GA<br>As           | p 1              | ATA<br>[le<br>[75 | G(         | ST<br>Y          | 528     |
| • | ATA<br>Ile        | CAA<br>Glr        | A AF         | M I              | AT 1<br>lis 1<br>.80 | TTT<br>Phe         | GCA<br>Ala        | TCA<br>Ser        | AGG<br>Arg           | GAA<br>Glu<br>185  | 1 G.           | GC :             | TTC<br>Phe       | GGG<br>Gly        | GGT<br>Gly        | AT.<br>11          | e M              | ATG<br>let        | CA<br>G1   | A<br>.n          | 576     |
|   | ATG<br>Met        | AAG<br>Lys        | Ph<br>19     | ie c             | ya I                 | CCA<br>Pro         | GAA<br>Glu        | TAT<br>Tyr        | GTA<br>Val<br>200    | AGC<br>Ser         | G:<br>Va       | TA 7             | Phe              | AAT<br>Asn        | AAT<br>Asn<br>205 | GT<br>Va           | T C              | AA<br>ln          | GA<br>G1   | A<br>u           | 624     |
|   | AAC<br>Asn        | AAA<br>Lys<br>210 | GG<br>G1     | y A              | CA A<br>la S         | GT .<br>Ser        | ATA<br>Ile        | TTT<br>Phe<br>215 | AAT<br>Asn           | AGA<br>Arg         | CC<br>Az       | GT (             | Sly ?            | TAT<br>Tyr<br>220 | TTT<br>Phe        | TC/<br>Sei         | A G              | AT<br>sp          | CC<br>Pr   | A<br>O           | 672     |
|   | GCC<br>Ala<br>225 | TTG<br>Leu        | AT.          | A T              | TA A<br>eu M         | ec i               | CAT<br>His<br>230 | GAA<br>Glu        | CTT<br>Leu           | ATA<br>Ile         | C#<br>Hi       | s V              | TT 1<br>al 1     | TTA<br>Leu        | CAT<br>His        | GG#<br>Gly         | A T              | eu                | TA'<br>Ty: | r                | 720     |
|   | GGC<br>Gly        | ATT<br>Ile        | AA.<br>Ly:   | A G              | TE A                 | AT (<br>sp )<br>45 | GAT<br>Asp        | TTA<br>Leu        | CCA<br>Pro           | ATT<br>Ile         | GI<br>Va<br>25 | l P              | CA A             | AAT<br>Asn        | GAA<br>Glu        | AAA<br>Lys         | L                | AA<br>ys<br>55    | TT:<br>Pho | r                | 768     |
|   | TTT .<br>Phe      | ATG<br>Met        | CA<br>Gl:    | A TO             | er T                 | CA (               | GAT<br>Asp        | GCT<br>Ala        | ATA<br>Ile           | CAG<br>Gln<br>265  | GC<br>Al       | A G<br>a G       | AA G<br>lu G     | BAA (             | CTA<br>Leu        | TAT<br>Tyr<br>270  | T                | CA '              | TT:<br>Phe | ;<br>;           | <br>916 |
|   |                   |                   |              |                  |                      |                    |                   |                   |                      |                    |                |                  |                  |                   |                   |                    |                  |                   |            |                  |         |



- 88 -

|   | •                 |                   |                       |                   |                   |                   |                    |                   |                   |                   |                   |                    |                   |                       |                   | •                 | •    |               |
|---|-------------------|-------------------|-----------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-----------------------|-------------------|-------------------|------|---------------|
|   | GGA<br>Gly        | GGA<br>Gly        | CAA<br>Gln<br>275     | GAT<br>Asp        | CCC<br>Pro        | AGC /<br>Ser      | Ile                | ATA<br>Ile<br>280 | ACT<br>Thr        | CCT<br>Pro        | TCT<br>Ser        | ACG<br>Thr         | GAT<br>Asp<br>285 | AAA<br>Lys            | AGT<br>Ser        | ATC<br>Ile        | 8    | 64            |
|   | TAT<br>Tyr        | GAT<br>Asp<br>290 | AAA<br>Lys            | GTT<br>Val        | TTG<br>Leu        | Gln .             | AAT<br>Asn<br>295  | TTT<br>Phe        | AGA<br>Arg        | GGG<br>Gly        | ATA<br>Ile        | GTT<br>Val<br>300  | GAT<br>Asp        | AGA<br>Arg            | CTT<br>Leu        | AAC<br>Asn        | 9    | 12            |
|   | AAG<br>Lys<br>305 | GTT<br>Val        | TTA<br>Leu            | GTT<br>Val        | TGC<br>Cys        | ATA<br>Ile<br>310 | TCA<br>Ser         | GAT<br>Asp        | CCT<br>Pro        | AAC<br>Asn        | ATT<br>Ile<br>315 | AAT<br>Asn         | ATT<br>Ile        | AAT<br>Asn            | ATA<br>Ile        | TAT<br>Tyr<br>320 | 9    | 60            |
|   | AAA<br>Lys        | AAT<br>Asn        | AAA<br>Lys            | TTT<br>Phe        | AAA<br>Lys<br>325 | GAT<br>Asp        | AAA<br>Lys         | TAT<br>Tyr        | AAA<br>Lys        | TTC<br>Phe<br>330 | GTT<br>Val        | GAA<br>Glu         | GAT<br>Asp        | TCT<br>Ser            | GAG<br>Glu<br>335 | GGA<br>Gly        | 10   | 800           |
|   | AAA<br>Lys        | TAT<br>Tyr        | AGT<br>Ser            | ATA<br>Ile<br>340 | Asp               | GTA<br>Val        | GAA<br>Glu         | Ser               | TTT<br>Phe<br>345 | GAT<br>Asp        | AAA<br>Lys        | TTA<br>Leu         | TAT               | AAA<br>Lys<br>350     | AGC<br>Ser        | TTA<br>Leu        | 10   | )56           |
|   | ATG<br>Met        | TTT<br>Phe        | GGT<br>Gly<br>355     | TTT<br>Phe        | ACA<br>Thr        | GAA<br>Glu        | ACT<br>Thr         | AAT<br>Asn<br>360 | ATA<br>Ile        | GCA<br>Ala        | GAA<br>Glu        | AAT                | TAT<br>Tyr<br>365 | AAA<br>Lys            | ATA<br>Ile        | AAA<br>Lys        | ïı   | 104           |
|   | ACT<br>Thr        | AGA<br>Arg<br>370 | Ala                   | TCT<br>Ser        | TAT<br>Tyr        | TTT<br>Phe        | AGT<br>Ser<br>375  | GAT<br>Asp        | TCC<br>Ser        | TTA<br>Leu        | Pro               | CCA<br>Pro<br>380  | GTA<br>Val        | AAA<br>Lys            | ATA<br>Ile        | AAA<br>Lys        | - 11 | 152           |
|   | AAT<br>Asn<br>385 | TTA<br>Leu        | TTA<br>Leu            | GAT<br>Asp        | AAT<br>Asn        | GAA<br>Glu<br>390 | ATC<br>Ile         | TAT<br>Tyr        | ACT<br>Thr        | ATA<br>Ile        | GAG<br>Glu<br>395 | GAA<br>Glu         | GGG<br>Gly        | TTT<br>Phe            | AAT<br>Asn        | ATA<br>Ile<br>400 | 12   | 200           |
|   | TCT<br>Ser        | GAT<br>Asp        | AAA<br>Lys            | GAT<br>Asp        | ATG<br>Met<br>405 | GAA<br>Glu        | AAA<br>Lys         | GAA<br>Glu        | TAT<br>Tyr        | AGA<br>Arg<br>410 | GGT<br>Gly        | CAG                | AAT<br>Asn        | AAA<br>Lys            | GCT<br>Ala<br>415 | ATA<br>Ile        | 12   | 248           |
| - | AAT<br>Asn        | AAA<br>Lys        | CAA<br>Gln            | GCT<br>Ala<br>420 | Tyr               | GAA<br>Glu        | GAA<br>Glu         | ATT               | AGC<br>Ser<br>425 | Lys               | GAG<br>Glu        | CAT                | TTG<br>Leu        | GCT<br>Ala<br>430     | GTA<br>Val        | TAT               | 1:   | 296           |
|   | AAG<br>Lys        | ATA<br>Ile        | CAA<br>Gln<br>435     | Met               | TGT<br>Cys        | AAA<br>Lys        | AGT<br>Ser         | GTT<br>Val<br>440 | Lys               | GCT<br>Ala        | CCA<br>Pro        | GGA<br>Gly         | ATA<br>Ile<br>445 | Cla                   | ATT<br>Ile        | GAT<br>Asp        | 1    | 344           |
|   | GTI<br>Val        | GAT<br>Asp        | Asn                   | GAA<br>Glu        | GAT<br>Asp        | TTG<br>Leu        | TTC<br>Phe<br>455  | TTT               | ATA<br>Ile        | GCT<br>Ala        | GAT<br>Asp        | AAA<br>Lys<br>460  | ASI               | AGT<br>Ser            | TTT<br>Phe        | TCA               | 1    | 392           |
|   | GAT<br>Asp        | Asp               | TTA<br>Lev            | TCT<br>Ser        | AAA<br>Lys        | AAC<br>Asn<br>470 | Glu                | AGA<br>Arg        | ATA<br>Ile        | GAA<br>Glu        | TAT<br>Tyr<br>475 | ABI                | ACA<br>Thr        | CAG<br>Glr            | AGT<br>Ser        | AAT<br>Asn<br>480 | . 1  | 440           |
|   | TAT               | T ATA             | GAZ<br>Glu            | LAA A<br>Laa Laa  | GAC<br>Asp<br>485 | Phe               | CCT<br>Pro         | ATA<br>Ile        | TAA A             | GAA<br>Glu<br>490 | r rer             | ATT                | r TTA             | A GAI                 | ACT<br>Thi<br>495 | GAT<br>Asp        | 1    | .488          |
|   | TT                | A ATI             | A AG                  | r AAA<br>r Lys    | ; Ile             | GAA<br>Glu        | TTA<br>Lev         | CCA               | A AGT             | GII               | A ÄAT<br>1 Asi    | ACA<br>Th:         | A GAI             | A TCA<br>1 Set<br>510 | L                 | ACT<br>Thr        |      | L <b>53</b> 6 |
|   | GA'               | r TT<br>p Ph      | r AA'<br>e As:<br>51: | n Vai             | A GAT<br>L Asp    | GTI<br>Val        | CCA<br>Pro         | GT/<br>Va.<br>52  | T TA              | GA)               | A AAI<br>Lys      | A CA               | n CC<br>n Pro     | J AI                  | r AT              | A AAA<br>e Lys    | :    | 1584          |
|   | AA<br>Ly          | A AT's Il         | e Ph                  | T AC              | A GAT             | GAA               | AAAA<br>AST<br>SSS | n Th              | C ATO             | C TT              | r CA<br>e Gl:     | A TA<br>n Ty<br>54 | I LE              | A TA                  | C TC<br>r Se      | r CAG             |      | 1632          |
|   |                   |                   |                       |                   |                   |                   |                    |                   |                   |                   |                   |                    |                   |                       |                   |                   |      |               |



- 89 -

|                               |                   |                    |                    |                      |                    |                   |                   |                      |                    | _                  | _                  |                    |                   |                    | •                    | •                   |
|-------------------------------|-------------------|--------------------|--------------------|----------------------|--------------------|-------------------|-------------------|----------------------|--------------------|--------------------|--------------------|--------------------|-------------------|--------------------|----------------------|---------------------|
| AC<br>Th<br>54                | I PA              | T CO               | T CI               | TA GA<br>eu As       | T AT<br>p Il<br>55 | e Arg             | A GA<br>G As      | T AT                 | A AG<br>e Se       | T TI<br>r Le<br>55 | u Th               | CA TO              | T TO              | CA T               | he As                | AT 1680<br>sp<br>60 |
| GA<br>As                      | T GC<br>p Al      | A TI<br>a Le       | A TI<br>u Le       | TA TT<br>tu Ph<br>56 | e Se               | T AA(<br>r Asi    | C AA              | A GT<br>s Va         | T TA<br>1 Ty<br>57 | r Se               | A TI<br>r Ph       | T TI<br>e Ph       | T TO              | er M               | TG GA<br>et As<br>75 | AT 1728<br>sp       |
| TA<br>Ty                      | T AT              | T AA<br>e Ly       | A AC<br>s Th<br>58 | T GC                 | T AA'<br>a Ası     | r AAA<br>n Lys    | A GTO             | G GT<br>l Val<br>589 | l Gl               | A GC<br>u Al       | A GG<br>a Gl       | A TT<br>y Le       | u Ph              | T GO               | CA GO<br>la Gl       | GT 1776<br>Ly       |
| TG                            | G GT<br>p Va      | G AA<br>l Ly<br>59 | s GI               | G ATA                | A GTA              | A AAI<br>l Asn    | GAT<br>Asp<br>600 | Phe                  | r GT               | A ATO              | C GA<br>e Gl       | A GC<br>u Al<br>60 | a As              | T AF               | VA AG                | SC 1824             |
| AA'<br>Asi                    | T AC              | r Me               | G GA'<br>t As      | T AAJ<br>p Lys       | A ATT              | GCA<br>Ala<br>615 | Asp               | ATA<br>Ile           | TC?                | r CT)              | A AT<br>110<br>620 | e Va               | T CC<br>1 Pr      | T TA               | T AT                 | TA 1872<br>e        |
| GGI<br>G1 <sub>3</sub><br>625 | Lei               | A GC'<br>1 Ala     | T TT               | A AAT<br>u Asr       | GTA<br>Val<br>630  | . Gly             | AAT<br>Asn        | GAA<br>Glu           | ACA<br>Thr         | GCT<br>Ala<br>635  | a Ly               | A GG               | A AA<br>Y As:     | T TI               | T GA<br>e Gl<br>64   | u                   |
| AA7<br>Asr                    | r GCT<br>n Ala    | TT.                | r GA(<br>≥ Glu     | G ATT                | : Ala              | GGA<br>Gly        | GCC<br>Ala        | AGT<br>Ser           | ATI<br>Ile<br>650  | : Leu              | CT)<br>Lei         | A GAI<br>1 Glu     | A TT              | T AT<br>e Il<br>65 | e Pro                | A 1968              |
| GAA<br>Glu                    | CTI<br>Lev        | TT!<br>Let         | ATA<br>1116<br>660 | A CCT<br>Pro         | GTA<br>Val         | GTT<br>Val        | GGA<br>Gly        | GCC<br>Ala<br>665    | Phe                | TTA<br>Leu         | TTA<br>Let         | A GA#              | TCI<br>Sei<br>670 | r Ty               | T AT                 | T 2016              |
| GAC<br>Asp                    | AAI<br>Asn        | Lys<br>675         | Asn                | AAA<br>Lys           | ATT                | ATT<br>Ile        | AAA<br>Lys<br>680 | ACA<br>Thr           | ATA<br>Ile         | GAT<br>Asp         | AAT<br>Asn         | GCT<br>Ala<br>689  | Lei               | A AC               | T AAJ<br>r Lys       | A 2064<br>5         |
| AGA                           | AAT<br>Asn<br>690 | GIU                | AAA<br>Lys         | TGG<br>Trp           | Ser                | GAT<br>Asp<br>695 | Met               | TAC<br>Tyr           | GGA<br>Gly         | Leu                | ATA<br>Ile<br>700  | : Val              | GCC<br>Ala        | G CA               | A TGC<br>n Trp       | 2112                |
| CTC<br>Leu<br>705             | TCA               | ÀCA<br>Thr         | GTT<br>Val         | AAT<br>Asn           | ACT<br>Thr<br>710  | CAA<br>Gln        | TTT<br>Phe        | TAT                  | ACA<br>Thr         | ATA<br>Ile<br>715  | AAA<br>Lys         | GAG<br>Glu         | GGA<br>Gly        | ATO<br>Met         | TATE TYPE            |                     |
| AAG<br>Lys                    | GCT<br>Ala        | TTA<br>Leu         | AAT<br>Asn         | TAT<br>Tyr<br>725    | CAA<br>Gln         | GCA<br>Ala        | CAA<br>Gln        | GCA<br>Ala           | TTG<br>Leu<br>730  | GAA<br>Glu         | GAA<br>Glu         | ATA<br>Ile         | ATA               | Lys<br>735         | Tyr                  | 2208                |
| AGA<br>Arg                    | TAT               | AAT<br>Asn         | ATA<br>Ile<br>740  | TAT<br>Tyr           | TCT<br>Ser         | GAA<br>Glu        | AAA<br>Lys        | GAA<br>Glu<br>745    | AAG<br>Lys         | TCA<br>Ser         | AAT<br>Asn         | ATT                | AAC<br>Asn<br>750 | Ile                | GAT<br>Asp           | 2256                |
| TTT<br>Phe                    | TAA<br>Asn        | GAT<br>Asp<br>755  | ATA<br>Ile         | AAT<br>Asn           | TCT<br>Ser         | Lys               | CTT<br>Leu<br>760 | AAT<br>Asn           | GAG<br>Glu         | GGT<br>Gly         | ATT                | AAC<br>Asn<br>765  | CAA<br>Gln        | GCT<br>Ala         | ATA                  | 2304                |
| GAT<br>Asp                    | AAT<br>Asn<br>770 | ATA<br>Ile         | AAT<br>Asn         | AAT<br>Asn           | TTT<br>Phe         | ATA<br>Ile<br>775 | TAA<br>Ran        | GGA<br>Gly           | TGT<br>Cys         | TCT<br>Ser         | GTA<br>Val<br>780  | TCA<br>Ser         | TAT<br>Tyr        | TTA<br>Leu         | ATG<br>Met           | 2352                |
| AAA<br>Lys<br>785             | AAA<br>Lys        | ATG<br>Met         | ATT<br>Ile         | CCA<br>Pro           | TTA<br>Leu<br>790  | GCT (<br>Ala      | GTA<br>Val        | GAA<br>Glu           | AAA<br>Lys         | TTA<br>Leu<br>795  | CTA<br>Leu         | GAC<br>Asp         | TTT<br>Phe        | GAT<br>Asp         | AAT<br>Asn<br>800    | 2400                |
| ACT<br>Thr                    | CTC<br>Leu        | AAA<br>Lys         | AAA<br>Lys         | AAT<br>Asn<br>805    | TTG<br>Leu         | TTA .<br>Leu .    | AAT<br>Asn        | Tyr                  | ATA<br>Ile<br>810  | GAT<br>Asp         | GAA<br>Glu         | AAT<br>Asn         | AAA<br>Lys        | TTA<br>Leu<br>815  | TAT<br>Tyr           | 2448                |



- 90 -

|                   | 1                      |                    |                       |                    |                    |                    |                    |                       |                    |                    |                    |                        |                    |                    |                    |            |           |
|-------------------|------------------------|--------------------|-----------------------|--------------------|--------------------|--------------------|--------------------|-----------------------|--------------------|--------------------|--------------------|------------------------|--------------------|--------------------|--------------------|------------|-----------|
| T7<br>Le          | rG AT                  | r GG<br>e Gly      | A AGI<br>Y Ser<br>820 | MIC                | A GAA<br>A Glu     | TAT<br>Tyr         | GAZ<br>Glu         | A AAA<br>1 Lys<br>825 | s Ser              | A AAA<br>C Lys     | A GTZ<br>Va:       | A AAT<br>l Asn         | AAA<br>Lys<br>830  | Ty                 | TTG<br>Leu         | 2          | 496       |
| AA<br>Ly          | A ACC                  | 2 ATT              | e met                 | CCG<br>Pro         | TTT<br>Phe         | GAT<br>Asp         | CTI<br>Leu<br>840  | Sex                   | ATA                | TAT                | ACC<br>Thi         | AAT<br>Asn<br>845      | Asp                | ACZ<br>Thi         | A ATA              |            | 544       |
| CT<br>Le          | 'A ATA<br>u Ile<br>850 | - GIL              | A ATG                 | TTI<br>Phe         | AAT<br>Asn         | AAA<br>Lys<br>855  | TYT                | AAT<br>Asn            | AGC<br>Ser         | GAA<br>Glu         | ATT                | : Leu                  | AAT<br>Asn         | AAT<br>Asn         | ATT                | 2:         | 592       |
| AT<br>11<br>86    | C TTA<br>e Leu<br>5    | AAT<br>ASN         | TTA<br>Leu            | AGA<br>Arg         | TAT<br>Tyr<br>870  | AAG<br>Lys         | GAT<br>Asp         | ' AAT<br>Asn          | AAT<br>Asn         | TTA<br>Leu<br>875  | ATA                | GAT<br>Asp             | TTA<br>Leu         | TCA<br>Ser         | GGA<br>Gly<br>880  | <b>2</b> ( | 640       |
| TA'<br>Ty         | T GGG<br>r Gly         | GCA<br>Ala         | AAG<br>Lys            | GTA<br>Val<br>885  | GAG<br>Glu         | GTA<br>Val         | TAT<br>Tyr         | GAT<br>Asp            | GGA<br>Gly<br>890  | GTC<br>Val         | <b>GA</b> G        | CTT                    | AAT<br>Asn         | GAT<br>Asp<br>895  | Lys                | 26         | 588       |
| AA'<br>Asi        | T CAA<br>n Gln         | TTT<br>Phe         | Lys<br>900            | TTA<br>Leu         | ACT<br>Thr         | AGT<br>Ser         | TCA<br>Ser         | GCA<br>Ala<br>905     | AAT<br>Asn         | AGT<br>Ser         | AAG<br>Lys         | ATT                    | AGA<br>Arg<br>910  | GTG<br>Val         | ACT<br>Thr         | 27         | 736       |
| CAI<br>Gl:        | A AAT<br>n Asn         | CAG<br>Gln<br>915  | AAT<br>Asn            | ATC<br>Ile         | ATA<br>Ile         | TTT<br>Phe         | AAT<br>Asn<br>920  | AGT<br>Ser            | GTG<br>Val         | TTC<br>Phe         | CTT<br>Léu         | GAT<br>Asp<br>925      | TTT<br>Phe         | AGC<br>Ser         | GTT<br>Val         | 27         | 784       |
| AG(               | TTT<br>Phe<br>930      | TGG<br>Trp         | ATA<br>Ile            | AGA<br>Arg         | ATA<br>Ile         | CCT<br>Pro<br>935  | AAA<br>Lys         | TAT                   | AAG<br>Lys         | AAT<br>Asn         | GAT<br>Asp<br>940  | GGT<br>Gly             | ATA<br>Ile         | CAA<br>Gln         | AAT<br>Asn         | 28         | 32        |
| TAT<br>Ty:        | T ATT                  | CAT<br>His         | AAT<br>Asn            | GAA<br>Glu         | TAT<br>Tyr<br>950  | ACA<br>Thr         | ATA<br>Ile         | ATT<br>Ile            | AAT<br>Asn         | TGT<br>Cys<br>955  | ATG<br>Met         | AAA<br>Lys             | AAT<br>Asn         | AAT<br>Asn         | TCG<br>Ser<br>960  | 28         | 80        |
| GGC<br>Gly        | TGG<br>Trp             | AAA<br>Lys         | ATA<br>Ile            | TCT<br>Ser<br>965  | ATT<br>Ile         | AGG<br>Arg         | GGT<br>Gly         | AAT<br>Asn            | AGG<br>Arg<br>970  | ATA<br>Ile         | ATA<br>Ile         | TGG<br>Trp             | ACT<br>Thr         | TTA<br>Leu<br>975  | ATT<br>Ile         | 29         | 28        |
| GAT<br>Asp        | T ATA                  | AAT<br>Asn         | GGA<br>Gly<br>980     | AAA<br>Lys         | ACC<br>Thr         | AAA<br>Lys         | TCG<br>Ser         | GTA<br>Val<br>985     | TTT<br>Phe         | TTT Phe            | GAA<br>Glu         | TAT<br>Tyr             | AAC<br>Asn<br>990  | ATA<br>Ile         | AGA<br>Arg         | 29         | 76        |
| GAA<br>Glu        | GAT<br>Asp             | ATA<br>Ile<br>995  | TCA<br>Ser            | GAG<br>Glu         | TAT<br>Tyr         | Ile                | AAT<br>Asn<br>1000 | Arg                   | TGG<br>Trp         | TTT<br>Phe         | TTT<br>Phe         | GTA<br>Val<br>1005     | Thr                | ATT<br>Ile         | ACT<br>Thr         | 30:        | 24        |
| AAT<br>Asn        | AAT<br>Asn<br>1010     | Leu                | AAT<br>Asn            | AAC<br>Asn         | Ala                | AAA<br>Lys<br>1015 | Ile                | TAT<br>Tyr            | ATT<br>Ile         | Asn                | GGT<br>Gly<br>1020 | Lys                    | CTA<br>Leu         | GAA<br>Glu         | TCA<br>Ser         | 30         | 72        |
| AAT<br>Asn<br>102 | ACA<br>Thr             | GAT<br>Asp         | ATT<br>Ile            | Lys                | GAT<br>Asp<br>1030 | ATA<br>Ile         | AGA<br>Arg         | GAA<br>Glu            | Val                | ATT<br>Ile<br>1035 | Ala                | TAA<br>neA             | GGT<br>Gly         | GAA<br>Glu         | ATA<br>Ile<br>1040 | 31:        | 20        |
| ATA<br>Ile        | TTT                    | AAA<br>Lys         | Leu .                 | GAT<br>Asp<br>1045 | Gly .              | GAT<br>Asp         | ATA<br>Ile         | Asp                   | AGA<br>Arg<br>1050 | Thr                | CAA<br>Gln         | TTT .<br>Phe           | Ile                | TGG<br>Trp<br>1055 | Met                | 310        | <b>68</b> |
| AAA<br>Lys        | TAT                    | TTC<br>Phe         | AGT<br>Ser<br>1060    | ATT<br>Ile         | TTT :<br>Phe :     | AAT .<br>Asn '     | Thr                | GAA<br>Glu<br>1065    | TTA<br>Leu         | AGT<br>Ser         | CAA<br>Gln         | Ser .                  | AAT<br>Asn<br>1070 | ATT<br>Ile         | GAA<br>Glu         | 323        | 16        |
| GAA<br>Glu        | AGA<br>Arg             | TAT<br>Tyr<br>1075 | Lys                   | ATT (              | CAA '              | Ser                | TAT<br>Tyr<br>1080 | Ser                   | GAA<br>Glu         | TAT<br>Tyr         | TTA<br>Leu         | AAA (<br>Lys /<br>1085 | GAT<br>Asp         | TTT<br>Phe         | TGG<br>Trp         | 326        | 54        |
|                   |                        |                    |                       |                    |                    |                    |                    |                       |                    |                    |                    |                        |                    |                    |                    |            |           |



- 91 -

| GG#<br>Gly        | AAT<br>Asn<br>109 | CCT<br>Pro<br>0    | TTA<br>Leu         | ATG<br>Met         | TAC                | AAT<br>Asn<br>109 | Lys                | GAA<br>Glu         | TAT                | TAT               | ATG<br>Met<br>110 | Phe                | AAT<br>Asn         | GCG<br>Ala         | GGG<br>Gly         | 3312 |
|-------------------|-------------------|--------------------|--------------------|--------------------|--------------------|-------------------|--------------------|--------------------|--------------------|-------------------|-------------------|--------------------|--------------------|--------------------|--------------------|------|
| AAT<br>Asn<br>110 | Lys               | AAT<br>Asn         | TCA<br>Ser         | TAT<br>Tyr         | ATT<br>Ile<br>1110 | Lys               | CTA<br>Leu         | AAG<br>Lys         | AAA<br>Lys         | GAT<br>Asp<br>111 | Ser               | CCT                | GTA<br>Val         | GGT<br>Gly         | GAA<br>Glu<br>1120 | 3360 |
| ATT               | TTA<br>Leu        | ACA<br>Thr         | CGT<br>Arg         | AGC<br>Ser<br>1125 | rys                | TAT<br>Tyr        | AAT<br>Asn         | CAA<br>Gln         | AAT<br>Asn<br>1130 | Ser               | AAA<br>Lys        | TAT<br>Tyr         | ATA<br>Ile         | AAT<br>Asn<br>1135 | Tyr                | 3408 |
| AGA<br>Arg        | GAT<br>Asp        | TTA<br>Leu         | TAT<br>Tyr<br>1140 | TTE                | GGA<br>Gly         | GAA<br>Glu        | AAA<br>Lys         | TTT<br>Phe<br>1145 | Ile                | ATA<br>Ile        | AGA<br>Arg        | AGA<br>Arg         | AAG<br>Lys<br>1150 | Ser                | AAT<br>Asn         | 3456 |
| TCT<br>Ser        | CAA<br>Gln        | TCT<br>Ser<br>1155 | TIE                | AAT<br>Asn         | GAT<br>Asp         | GAT<br>Asp        | ATA<br>Ile<br>1160 | Val                | AGA<br>Arg         | AAA<br>Lys        | GAA<br>Glu        | GAT<br>Asp<br>1165 | Tyr                | ATA<br>Ile         | TAT<br>Tyr         | 3504 |
| CTA<br>Leu        | GA                |                    |                    |                    |                    |                   |                    | •                  |                    |                   | •                 |                    |                    |                    |                    | 3509 |

#### (2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1169 amino acids(B) TYPE: amino acid

  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Met Pro Val Thr Ile Asn Asn Phe Asn Tyr Asn Asp Pro Ile Asp Asn

Asn Asn Ile Ile Met Met Glu Pro Pro Phe Ala Arg Gly Thr Gly Arg

Tyr Tyr Lys Ala Phe Lys Ile Thr Asp Arg Ile Trp Ile Ile Pro Glu

Arg Tyr Thr Phe Gly Tyr Lys Pro Glu Asp Phe Asn Lys Ser Ser Gly 50 60

Ile Phe Asn Arg Asp Val Cys Glu Tyr Tyr Asp Pro Asp Tyr Leu Asn 65 70 75 80

Thr Asn Asp Lys Lys Asn Ile Phe Leu Gln Thr Met Ile Lys Leu Phe

Asn Arg Ile Lys Ser Lys Pro Leu Gly Glu Lys Leu Leu Glu Met Ile 100 105 110

Ile Asn Gly Ile Pro Tyr Leu Gly Asp Arg Arg Val Pro Leu Glu Glu

Phe Asn Thr Asn Ile Ala Ser Val Thr Val Asn Lys Leu Ile Ser Asn

Pro Gly Glu Val Glu Arg Lys Lys Gly Ile Phe Ala Asn Leu Ile Ile 155

Phe Gly Pro Gly Pro Val Leu Asn Glu Asn Glu Thr Ile Asp Ile Gly

Ile Gln Asn His Phe Ala Ser Arg Glu Gly Phe Gly Gly Ile Met Gln Met Lys Phe Cys Pro Glu Tyr Val Ser Val Phe Asn Asn Val Gln Glu 200 Asn Lys Gly Ala Ser Ile Phe Asn Arg Arg Gly Tyr Phe Ser Asp Pro Ala Leu Ile Leu Met His Glu Leu Ile His Val Leu His Gly Leu Tyr Gly Ile Lys Val Asp Asp Leu Pro Ile Val Pro Asn Glu Lys Lys Phe Phe Met Gln Ser Thr Asp Ala Ile Gln Ala Glu Glu Leu Tyr Thr Phe 265 Gly Gly Gln Asp Pro Ser Ile Ile Thr Pro Ser Thr Asp Lys Ser Ile Tyr Asp Lys Val Leu Gln Asn Phe Arg Gly Ile Val Asp Arg Leu Asn Lys Val Leu Val Cys Ile Ser Asp Pro Asn Ile Asn Ile Asn Ile Tyr 315 Lys Asn Lys Phe Lys Asp Lys Tyr Lys Phe Val Glu Asp Ser Glu Gly Lys Tyr Ser Ile Asp Val Glu Ser Phe Asp Lys Leu Tyr Lys Ser Leu 340 Met Phe Gly Phe Thr Glu Thr Asn Ile Ala Glu Asn Tyr Lys Ile Lys Thr Arg Ala Ser Tyr Phe Ser Asp Ser Leu Pro Pro Val Lys Ile Lys 375 Asn Leu Leu Asp Asn Glu Ile Tyr Thr Ile Glu Glu Gly Phe Asn Ile 390 Ser Asp Lys Asp Met Glu Lys Glu Tyr Arg Gly Gln Asn Lys Ala Ile 405 Asn Lys Gln Ala Tyr Glu Glu Ile Ser Lys Glu His Leu Ala Val Tyr Lys Ile Gln Met Cys Lys Ser Val Lys Ala Pro Gly Ile Cys Ile Asp 440 Val Asp Asn Glu Asp Leu Phe Phe Ile Ala Asp Lys Asn Ser Phe Ser 455 Asp Asp Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn Tyr Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 490 Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asn Thr Glu Ser Leu Thr 505 Asp Phe Asn Val Asp Val Pro Val Tyr Glu Lys Gln Pro Ala Ile Lys

Lys Ile Phe Thr Asp Glu Asn Thr Il Phe Gln Tyr Leu Tyr Ser Gln Thr Phe Pro Leu Asp Ile Arg Asp Ile Ser Leu Thr Ser Ser Phe Asp 550 Asp Ala Leu Leu Phe Ser Asn Lys Val Tyr Ser Phe Phe Ser Met Asp Tyr Ile Lys Thr Ala Asn Lys Val Val Glu Ala Gly Leu Phe Ala Gly Trp Val Lys Gln Ile Val Asn Asp Phe Val Ile Glu Ala Asn Lys Ser Asn Thr Met Asp Lys Ile Ala Asp Ile Ser Leu Ile Val Pro Tyr Ile Gly Leu Ala Leu Asn Val Gly Asn Glu Thr Ala Lys Gly Asn Phe Glu Asn Ala Phe Glu Ile Ala Gly Ala Ser Ile Leu Leu Glu Phe Ile Pro Glu Leu Leu Ile Pro Val Val Gly Ala Phe Leu Leu Glu Ser Tyr Ile Asp Asn Lys Asn Lys Ile Ile Lys Thr Ile Asp Asn Ala Leu Thr Lys Arg Asn Glu Lys Trp Ser Asp Met Tyr Gly Leu Ile Val Ala Gln Trp Leu Ser Thr Val Asn Thr Gln Phe Tyr Thr Ile Lys Glu Gly Met Tyr Lys Ala Leu Asn Tyr Gln Ala Gln Ala Leu Glu Glu Ile Ile Lys Tyr Arg Tyr Asn Ile Tyr Ser Glu Lys Glu Lys Ser Asn Ile Asn Ile Asp Phe Asn Asp Ile Asn Ser Lys Leu Asn Glu Gly Ile Asn Gln Ala Ile Asp Asn Ile Asn Asn Phe Ile Asn Gly Cys Ser Val Ser Tyr Leu Met Lys Lys Met Ile Pro Leu Ala Val Glu Lys Leu Leu Asp Phe Asp Asn Thr Leu Lys Lys Asn Leu Leu Asn Tyr Ile Asp Glu Asn Lys Leu Tyr Leu Ile Gly Ser Ala Glu Tyr Glu Lys Ser Lys Val Asn Lys Tyr Leu Lys Thr Ile Met Pro Phe Asp Leu Ser Ile Tyr Thr Asn Asp Thr Ile Leu Ile Glu Met Phe Asn Lys Tyr Asn Ser Glu Ile Leu Asn Asn Ile Ile Leu Asn Leu Arg Tyr Lys Asp Asn Asn Leu Ile Asp Leu Ser Gly

#### - 94 -

Tyr Gly Ala Lys Val Glu Val Tyr Asp Gly Val Glu Leu Asn Asp Lys

Asn Gln Phe Lys Leu Thr Ser Ser Ala Asn Ser Lys Ile Arg Val Thr

Gln Asn Gln Asn Ile Ile Phe Asn Ser Val Phe Leu Asp Phe Ser Val 915 920 925

Ser Phe Trp Ile Arg Ile Pro Lys Tyr Lys Asn Asp Gly Ile Gln Asn 930 935

Tyr Ile His Asn Glu Tyr Thr Ile Ile Asn Cys Met Lys Asn Asn Ser

Gly Trp Lys Ile Ser Ile Arg Gly Asn Arg Ile Ile Trp Thr Leu Ile 965 970 975

Asp Ile Asn Gly Lys Thr Lys Ser Val Phe Phe Glu Tyr Asn Ile Arg

Glu Asp Ile Ser Glu Tyr Ile Asn Arg Trp Phe Phe Val Thr Ile Thr 995 1000 1005

Asn Asn Leu Asn Asn Ala Lys Ile Tyr Ile Asn Gly Lys Leu Glu Ser 1010 1015 1020

Asn Thr Asp Ile Lys Asp Ile Arg Glu Val Ile Ala Asn Gly Glu Ile 1025 1030 1035 1040

Ile Phe Lys Leu Asp Gly Asp Ile Asp Arg Thr Gln Phe Ile Trp Met 1045 1050 1055

Lys Tyr Phe Ser Ile Phe Asn Thr Glu Leu Ser Gln Ser Asn Ile Glu 1060 1065 1070

Glu Arg Tyr Lys Ile Gln Ser Tyr Ser Glu Tyr Leu Lys Asp Phe Trp 1075 1080 1085

Gly Asn Pro Leu Met Tyr Asn Lys Glu Tyr Tyr Met Phe Asn Ala Gly 1090 1095 1100

Asn Lys Asn Ser Tyr Ile Lys Leu Lys Lys Asp Ser Pro Val Gly Glu 1110 1115 1120

Ile Leu Thr Arg Ser Lys Tyr Asn Gln Asn Ser Lys Tyr Ile Asn Tyr 1125 1130 1135

Arg Asp Leu Tyr Ile Gly Glu Lys Phe Ile Ile Arg Arg Lys Ser Asn 1140 1145 1150

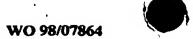
Ser Gln Ser Ile Asn Asp Asp Ile Val Arg Lys Glu Asp Tyr Ile Tyr 1155 1160 1165

Leu

## (2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2574 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

PCT/GB97/02273



- 95 -

## (ix) FEATURE:

(A) NAME/KEY: CDS (B) LOCATION:1..2574

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

|                   |                      |                   | -                 |                   |                       |                  |                   |                    |                |                       |                            |                   |                   |            |                      |      |
|-------------------|----------------------|-------------------|-------------------|-------------------|-----------------------|------------------|-------------------|--------------------|----------------|-----------------------|----------------------------|-------------------|-------------------|------------|----------------------|------|
| AT<br>Me          | G CC<br>t Pr<br>1    | A GI              | TT AC             | A AT              | 'A AA'<br>e Ası<br>5  | r Aar<br>n Ası   | r TT<br>n Ph      | T AA'<br>e Asi     | T TA'<br>n Ty: | r Aar                 | r Ga'<br>n As <sub>l</sub> | T CC              | T AT              | e As       | T AAT<br>p Asr<br>5  | r 46 |
| AA<br>As          | T AA<br>n As         | T AT              | e II              | T AT<br>e Me<br>0 | G ATO                 | G GAO            | G CC              | T CCI<br>Pro<br>25 | o Phe          | GCC<br>Ala            | AGA                        | A GG1             | r AC              | r Gl       | G AGA<br>y Arg       | 96   |
| Ty.               | r ly                 | 3                 | 5 AI              | a Pn              | e ras                 | ; ITE            | 4 (               | c Asr              | Arg            | , Ile                 | Trį                        | 11e<br>45         | : Ile             | Pr         | G GAA<br>o Glu       |      |
| AG                | A TA'<br>g Ty:<br>50 | LIN               | T TT<br>r Ph      | T GG/<br>e Gly    | A TAI<br>Y Tyr        | AAA<br>Lys<br>55 | Pro               | GAG<br>Glu         | GAT<br>Asp     | TTT<br>Phe            | AAT<br>Asn<br>60           | Lys               | AGT<br>Ser        | TC(        | C GGT<br>r Gly       | 192  |
| 69                | 5 Pile               | e Asi             | n Arg             | g Ast             | 70                    | Cys              | GIU               | Tyr                | Tyr            | 75                    | Pro                        | Asp               | Tyr               | Lei        | TAA A<br>nsA 1<br>08 |      |
| 1111              | . ASI                | ı Asj             | b PA              | 85<br>85          | asn<br>;              | ite              | Pne               | Leu                | 90             | Thr                   | Met                        | Ile               | Lys               | Let<br>95  |                      | 288  |
| ASII              | Arg                  | 1 116             | 100               | sser              | Lys                   | Pro              | Leu               | 105                | Glu            | Lys                   | Leu                        | Leu               | Glu<br>110        | Met        | ATT<br>Ile           | 336  |
| 116               | ASII                 | 115               | i                 | Pro               | 1yr                   | ren              | 120               | Asp                | Arg            | Arg                   | Val                        | Pro<br>125        | Leu               | Glu        |                      | 384  |
| PHE               | 130                  | · Ini             | ASN               | ııe               | GCT<br>Ala            | 135              | val               | Thr                | Val            | Asn                   | Lys<br>140                 | Leu               | Ile               | Ser        | Asn                  | 432  |
| 145               | GIA                  | GIU               | vai               | GIU               |                       | rys              | ràs               | GIÀ                | Ile            | Phe<br>155            | Ala                        | Asn               | Leu               | Ile        | Ile<br>160           | 480  |
| PHE               | GIY                  | PIO               | GIY               | 165               | GTT<br>Val            | ren              | Asn               | Glu                | 170            | Glu                   | Thr                        | Ile               | Asp               | Ile<br>175 | Gly                  | 528  |
| ATA<br>Ile        | CAA<br>Gln           | AAT<br>Asn        | CAT<br>His<br>180 | TTT               | GCA<br>Ala            | TCA<br>Ser       | AGG<br>Arg        | GAA<br>Glu<br>185  | Gly            | TTC Phe               | GGG<br>Gly                 | GGT<br>Gly        | ATA<br>Ile<br>190 | ATG<br>Met | CAA<br>Gln           | 576  |
| ATG<br>Met        | AAG<br>Lys           | TTT<br>Phe<br>195 | TGC               | CCA<br>Pro        | GAA<br>Glu            | Tyr              | GTA<br>Val<br>200 | AGC<br>Ser         | GTA<br>Val     | TTT /<br>Phe /        | Asn                        | AAT<br>Asn<br>205 | GTT<br>Val        | CAA<br>Gln | GAA<br>Glu           | 624  |
| Asn               | Lys<br>210           | GIÀ               | Ala               | ser               |                       | Phe .<br>215     | Asn .             | Arg .              | Arg (          | Gly :                 | Tyr<br>220                 | Phe :             | Ser .             | Asp        | Pro                  | 672  |
| GCC<br>Ala<br>225 | TTG<br>Leu           | ATA<br>Ile        | TTA<br>Leu        | Met               | CAT (<br>His (<br>230 | GAA (<br>Glu )   | CTT .<br>Leu      | ATA (              | His '          | GTT 7<br>Val 1<br>235 | FTA (<br>Leu )             | CAT (             | GGA<br>Gly        | Leu        | TAT<br>Tyr<br>240    | 720  |

# - 96 -

| TIT ATG CAA TCT ACA GAT GCT ATA CAG GCA GAA GAA CTA TAT ACA ITT Phe Met GIn Ser Thr Asp Als 11e Gin als Glu Glu Leu Tyr Thr Phe 260  GGA GGA CAA GAT CCC AGC ATC ATA ACT CCT TCT ACG GAT AAA AGT ATA CAG ITT CAG GIY GIY GIN Asp Pro Ser Ile 11e Thr Pro Ser Thr Asp Lys Ser Ile 275  TAT GAT AAA GTT TTG CAA AAT TTT AGA GGG ATA GTT GAT ACA CTT AAC TYR Asp Lys Val Leu Gin Asn Phe Arg Gly Ile Val Asp Arg Leu Asn 290  AAG GTT TTA GTT TGC ATA TCA GAT CCT AAC ATT AAT ATA ATA TAT TAT Lys Val Leu Val Cys Ile Ser Asp Pro Asn 11e Asn Ile Asn Ile Tyr 305  AAA AAT AAA TTT AAA GAT AAA TAT AAA TTC GTT GAA GAT TCA GAG GGG ATA CAT AAC Lys Asn Lys Phe Lys Asp Lys Tyr Lys Phe Val Glu Asp Ser Glu Gly 310  AAA AAT AAA TTT AAA GAT AAA TAT AAA TTC GTT GAA GAT TCT GAG GGG ALY GTT GAT ACT ATA ACT ACT   |   | ~                 |                    |                   |                    |                |              |                |                |                   |                |                |                |                |                |              |                | •    |
|--|---|-------------------|--------------------|-------------------|--------------------|----------------|--------------|----------------|----------------|-------------------|----------------|----------------|----------------|----------------|----------------|--------------|----------------|------|
| GGA GGA GAA GAT CCC AGC ATC ATA ACT CCT TCT ACG GAT AAA ACT ATC CGI GIY GIY GIN ASP PTO SET IIE IIE THY PTO SET THY ASP LYS SET IIE 280 TAT GAT AAA GTT TTG CAA AAT TTT AGA GGG ATA AT GAT GAT AGA CTT AAC 290 295  TAT GAT AAA GTT TTG CAA AAT TTT AGA GGG ATA GTT GAT AGA CTT AAC 290 295  AAG GTT TTA GTT TGC AAT ACT ACT ACC ACT AAT ATA TAT ATA TAT LYS 290 295  AAG GTT TTA GTT TGC AAT ACT ACT ACT AAT AAT ATA TAT ATA CLYS ASA LYS ASA |   |                   |                    |                   |                    | 2              | 45           | up 1           | cu P.          | 10 1.             | 25             | 0<br>T br      | O As           | n Gl           | u Ly           | s Ly<br>25   | s Phe          |      |
| TAT GAT AAA GTT TTG CAA AAT TTT AGA GGG ATA GTT GAT AAA GTT TAAC 290  AAG GTT TTA GTT AGA CATA TCA GAT CCT AAC ATT AAT ATA ATA TAT Lyx spa Lyx Val Leu Gin Aan Phe Arg Giy Ile Val Aap Arg Leu Aan 290  AAG GTT TTA GTT TGC ATA TCA GAT CCT AAC ATT AAT ATT AAT ATA TAT Lyx Val Leu Val Cyx Ile Ser Aap Pro Aan Ile Aan Ile Aan Ile Tyx 305  AAA AAT AAA TTT AAA GAT AAA TAT AAA TTT GAT GA  | • | ' TT              | T A                | IG C              |                    |                | CA G<br>hr A | AT G<br>sp A   | CT A:          | re G              | ru wr          | A GA<br>a Gli  | A GA<br>u Gl   | A CT.<br>u Le  | u Ty           | r Th         | A TTT<br>r Phe | 816  |
| AGG GTT TTA GTT TOC ATA TCA GAT CCT AAC ATT AAT ATT AAT ATA TAT LYS Val Leu Val Cys Ile Ser Asp Pro Asn Ile Asn Ile Asn Ile Tyr 305  AAA AAT AAA TTT AAA GAT AAA TAT AAA TTC GTT GAA GAT CTT GAG GGA 106 Lys Asn Lys Phe Lys Asp Lys Tyr Lys Phe Val Glu Asp Ser Glu Gly 325  AAA TAT AAT ATT GTT GTT GTA GAT GTA GAA GTT TTT GAT AAA ATT AAA AGC TTA Lys Tyr Ser Ile Asp Val Glu Ser Phe Asp Lys Leu Tyr Lys Ser Leu 340  ATG TTT GGT TTT ACA GAA ACT AAT ATA GCA GAA AAT TAT AAA ATA AAA ATA AAA TAT AGT ATA TTT ACA GAA ACT AAT ATA GAA GAA ATT TYS Lys Ile Lys 355  ACT AGA GCT TCT TAT TTT AGT GAT TCC TTA CCA CCA GTA AAA ATA AAA ATA AAA TAT ARA ASP ALS SER Leu Asp Asn Glu Ile Tyr Thr Ile Glu Glu Gly Phe Asn Ile 375  AAT TTA TTA GAT AAT GAA ATC TAT ACT ATA GAG GAA GGG TTT AAT ATA ASN Leu Leu Asp Asn Glu Ile Tyr Thr Ile Glu Glu Gly Phe Asn Ile 385  390  TCT GAT AAA GAT ATG GAA AAA GAA TAT AGC GAG GAT AAA GCT TAT 380  ACT AAA GAT ATG GAA AAA GAA TAT AGC AGG GAT AAA GCT TAT 380  TCT GAT AAA GAT ATG GAA AAA GAA TAT AGC AGG GAT AAA GCT TAT 380  TCT GAT AAA GAT ATG GAA AAA GAA TAT AGC AGG GAT CAG AAT AAA GCT TAT 380  AAT AAA CAA GCT TAT GAA GAA ATA GAA ATA AGA GTT ATA 380  AAT AAA CAA GCT TAT GAA GAA ATA AGC AAG GAG CAT TTG GCT GTA TAT 380  AAT AAA CAA GCT TAT GAA GAA ATT AGC AAG GAG CAT TTG GCT GTA TAT 380  AAT AAA CAA GCT TAT GAA GAA ATT AGC AAG GAG CAT TTG GCT GTA TAT 380  AAG ATA CAA ATG TGT AAA AGT GTT AAA GTT TAAA GCT AAA ATA GCT ATA 380  AAG ATA CAA ATG TAT GAA GAA ATT AGC AAG GAC CAT TTG GCT GTA TAT 380  AAG ATA CAA ATG TAT AAA AGT GTT AAA AGT TTA AAA GCT TAT AAA AGT TTT TCA 380  AAG ATA CAA ATG GAA AAA GT GTT AAA AGT GTT AAA AGT TTT TCA 435  GTT GAT AAA GAA TTG TCT TTC TTT ATA GCT GAA AAA AAT AGT TTT TCA 435  GTT GAT AAA GAA ATG GTT AAA AGT GTT AAA AGT TTT TCA 445  GTT GAT AAT GAA AAT GAA TTG TCT TTT ATA GCT GAA AAA AAT ACA CAG AGT AAT 380  AAG ATA CAA ATG GAA AAT GAA ATG GTT AAA GTT TTA TAT GAT ATT AAT AA   |   | GG<br>G1          | A GO<br>y Gl       | SA C<br>Ly G<br>2 | AA G<br>ln A<br>75 | AT C           | CC A         | GC A'<br>er I  | re 11          | re II             | T CC           | T TC           | r ACC          | r Ası          | p Ly           | A AG<br>s Se | T ATC          | 864  |
| AAA AAT AAA TIT AAA GAT AAA TAT AAA TIC GIT GAA GAT TCT GAG GGA Lys Asn Lys Phe Lys Asp Lys Tyr Lys Phe Val Glu Asp Ser Glu Gly 325 325 330  AAA TAT AGT ATA GAT GTA GAA AGT TIT GAT AAA TTA TAT AAA AGC TTA Lys Tyr Ser Ile Asp Val Glu Ser Phe Asp Lys Leu Tyr Lys Ser Leu 340  ATG TIT GGT TIT ACA GAA ACT AAT ATA GCA GAA AAT TAT AAA ATA AAA MET Phe Gly Phe Thr Glu Thr Asn Ile Ala Glu Asn Tyr Lys Ile Lys 355 365  ACT AGA GCT TCT TAT TIT AGT GAT TCC TTA CCA CCA GTA AAA ATA AAA THA AAA THA AGA GCT TTA ATA AGA GCT TCT TAT TIT AGT GAT TCC TTA CCA CCA GTA AAA ATA AAA THA AGA ST ATA AAA ATA AGA GCT TCT TAT TTA GAT ATA TATA GAA GGA GGG TTT AAT ATA ASN Leu Leu Asp Asn Glu Ile Tyr Thr Ile Glu Glu Gly Phe Asn Ile 395  TCT GAT AAA GAT ATG GAA AAA GAA TAT AGA GAG GGT CAG AAT AAA GCT ATA ASN Leu Leu Asp Met Glu Lys Glu Tyr Arg Gly Gln Asn Lys Ala Ile 405  AAT AAA CAA GCT TAT GAA GAA AAT AGC AAG GAG CAT TTG GCT GTA TAT ASN Lys Gln Ala Tlyr Glu Glu Glu Glu His Leu Ala Val Tyr 420  AAG ATA CAA ATG TAT GAA GAA ATT AGC AAG GAG CAT TTG GCT GTA TAT ASN Lys Gln Ala Tlyr Glu Glu Glu Ser Lys Glu His Leu Ala Val Tyr 420  AAG ATA CAA ATG TGT AAA AGT GTT AAA GCT CCA GGA ATA TAT TGT ATT GAT ASN Lys Ala Pro Gly His Cys Ile Asp 435  GTT GAT AAA GGT TAT GAA GAA GTT ATA GCT GAT AAA ATT TTT TCA 425  AAG ATA CAA ATG TGT AAA AGT GTT AAA GCT CCA GGA ATA TGT ATT GAT ATT GAT ASN Lys Ala Pro Gly His Cys Ile Asp 435  GTT GAT AAA GAA GTT TTG TTC TTT ATA GCT GAT AAA AAT AGT TTT TCA 455  GAT GAT TAT TCT AAA AAC GAA AGA ATA GAA TAT AAT ACA CAG AGT AAT AAA AAT AGT ATA GAA AAT GAA TTA TT  |   | TA<br>Ty          | T GA<br>r As<br>29 | T A               | AA G<br>ys V       | TT T           | rg C         | an and         | III EI         | T AG              | A GGG          | G ATA          | ≥ Val          | L Asp          | r Ag           | A CT         | T AAC<br>u Asn | 912  |
| AAA AAT AAA TTT AAA GAT AAA TAT AAA TTC GTT GAA GAT TCT GAG GGA Lys Asn Lys Phe Lys Asp Lys Tyr Lys Phe Val Glu Asp Ser Glu Gly 325  AAA TAT AGT ATA GAT GTA GAA AGT TTT GAT AAA TTA TAT AAA AGC TTA Lys Tyr Ser Ile Asp Val Glu Ser Phe Asp Lys Leu Tyr Lys Ser Leu 340  ATG TTT GGT TTT ACA GAA ACT AAT ATA GCA GAA AAT TAT AAA ATA AAA Met Phe Gly Phe Thr Glu Thr Asn Ile Ala Glu Asn Tyr Lys Ile Lys 355  ACT AGA GCT TCT TAT TTT AGT GAT TCC TTA CCA CCA GTA AAA ATA AAA Met Phe Gly Phe Thr Glu Thr Asn Ile Ala Glu Asn Tyr Lys Ile Lys 370  ACT AGA GCT TCT TAT TTT AGT GAT TCC TTA CCA CCA GTA AAA ATA AAA STA TATA TATA GAT AAT GAA ATC TAT ACT ATA GAG GAA GGG TTT AAT ATA ASN Leu Leu Asp Asn Glu Ile Tyr Thr Ile Glu Glu Gly Phe Asn Ile 385  TCT GAT AAA GAT ATG GAA AAA GAA TAT AGA GAG GAG   |   | AA<br>Ly<br>30    | G GI<br>s Va<br>5  | T T:<br>1 Le      | ra Gʻ              | TT TO<br>al Cy |              |                | A GA<br>r As   | T CC<br>p Pr      | T AA(<br>O Asi | ) ITE          | Asn            | ATT            | T AAT<br>S Asi | T ATA        | e Tyr          | 960  |
| ATG TTT GGT TTT ACA GAA ACT AAT ATA GCA GAA AAT TAT AAA ATA AAA TTA AGA GCT TAT TTT AGA GAA ACT TAT ACC ACA GTA AAA ATA AAA ATA AAA ATA ACA GAA ATA CAA ACT CAT ACT ACT ACT ACT ACT A  |   | AA<br>Ly:         | A AA<br>5 As       | T AA              | A Ti               | ,              |              | AT AA<br>Sp Ly | A TA<br>s Ty   | T AA<br>r Ly      | s Pne          | : val          | GAA<br>Glu     | GAT<br>Asp     | TC1            | Gli          | G GGA          | 1008 |
| ATG TTT GGT TTT ACA GAA ACT AAT ATA GCA GAA AAT TAT AAA ATA AAA  Met Phe Gly Phe Thr Glu Thr Asn Ile Ala Glu Asn Tyr Lys Ile Lys 365  ACT AGA GCT TCT TAT TTT ACT GAT TCC TTA CCA CCA GTA AAA ATA AAA  TTRA ATG Ala Ser Tyr Phe Ser Asp Ser Leu Pro Pro 380  AAT TTA TTA GAT AAT GAA ATC TAT ACT ATA GAG GAA GGG TTT AAT ATA Asn Leu Leu Asp Asn Glu Ile Tyr Thr Ile Glu Glu Gly Phe Asn Ile 395  TCT GAT AAA GAT ATG GAA AAA GAA TAT AGG GGT CAG AAT AAA GCT ATA Ser Asp Lys Asp Met Glu Lys Glu Tyr Arg Gly Gln Asn Lys Ala Ile 405  AAT AAA CAA GCT TAT GAA GAA ATT AGC AAG GAG CAT TTG GCT GTA TAT ASN Lys Gln Ala Tyr Glu Glu Ile Ser Lys Glu His Leu Ala Val Tyr 420  AAG ATA CAA ATG TGT AAA AGT GTT AAA GCT CCA GGA ATA TGT ATT GAT Lys Ile Gln Met Cys Lys Ser Val Lys Ala Pro Gly Ile Cys Ile Asp 445  GTT GAT AAT GAA GAT TTG TTC TTT ATA GCT GAT AAA AAT ACT TTT TCA ASP ASP ASP Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn 450  GAT GAT TTA TCT AAA AAC GAA AGA ATA GAA TAA AAT ACA CAG AGT AAT ASP ASP Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn 465  TATA ATA GAA AAT GAC TTC CCT ATA AAT GAA TAA AAT ACT GAT ACT ASP ASP Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn 470  TATA ATA GAA AAT GAC TTC CCT ATA AAT GAA TTA ATT ACA CAG AGT AAT Tyr Ile Glu Asn Asp Phe Pro Ile Asn Glu Tyr Asn Thr Gln Ser Asn 480  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT 1486  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA CAG AT TACT ATA AGT AAA AAT GAA TTA CCA AGT GAA AAT ACA CAG AT TACT ATA AGT AAA AAT GAA TTA CCA AGT GAA AAT ACA CAG ATT ACT ATA AGT AAA AAT GAA TTA CCA AGT GAA ATA ATA ACA CAG AGT AAT AATA AAA AAT AAA AAT GAA TTA CCA AGT GAA ATA ACT ACT ACT ATA ATA AGT AAA AAT GAA TTA CCA AGT GAA AAT ACA CAG AT TACT TAC  |   | AAI<br>Lys        | TA'                | T AG              |                    |                | T GI         | A GA<br>1 G1   | A AG'<br>u Se: | r Pue             | e Asp          | AAA<br>Lys     | TTA<br>Leu     | TAT            | Lys            | AGC<br>Ser   |                | 1056 |
| ACT AGA GCT TCT TAT TTT AGT GAT TCC TTA CCA CCA GTA AAA ATA AAA  The Arg ala Ser Tyr Phe Ser Asp Ser Leu Pro Pro Val Lys Ile Lys 380  AAT TTA TTA GAT AAT GAA ATC TAT ACT ATA GAG GAA GGG TTT AAT ATA ASP ASP ASP Glu Ile Tyr Thr Ile Glu Gly Phe Asm Ile 400  TCT GAT AAA GAT ATG GAA AAA GAA TAT AGA GGT CAG AAT AAA GCT ATA ASP Lys Asp Met Glu Lys Glu Tyr Arg Gly Gln Asm Lys Ala Ile 415  AAT AAA CAA GCT TAT GAA GAA ATT AGC AAG GAG CAT TTG GCT GTA TAT ASP Lys Glu Ala Tyr Glu Glu Ile Ser Lys Glu His Leu Ala Val Tyr A30  AAG ATA CAA ATG TGT AAA AGT GTT AAA GCT CCA GGA ATA TGT ATT GAT A35 Lys Ile Gln Met Cys Lys Ser Val Lys Ala Pro Gly Ile Cys Ile Asp 445  GTT GAT AAT GAA GAT TTG TC TTT ATA GCT GAT AAA AAT AGT TTT TCA 450  GAT GAT TTA TCT AAA AAC GAA AGA ATA GAA TAT AAT ACA CAG AGT AAT ASP ASP Leu Phe Phe Ile Ala Asp Lys Asm Ser Phe Ser 470  GAT GAT TTA TCT AAA AAC GAA AGA ATA GAA TAT AAT ACA CAG AGT AAT ASP ASP Leu Ser Lys Asm Glu Arg Ile Glu Tyr Asm Thr Gln Ser Asm 480  TAT ATA GAA AAT GAC TC CCT ATA AAT GAA TTA ATT TTA GAT ACT GAT THA ASP ASP Phe Pro Ile Asm Glu Leu Ile Leu Asp Thr Asp 485  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT 1536  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asm Thr Glu Ser Leu Thr  |   | ATC<br>Met        | TT'                |                   | <i>_</i>           | T AC           | A GA<br>r Gl | A AC           | r war          | T. TTE            | A GCA          | GAA<br>Glu     | AAT<br>Asn     | Tyr            |                |              | AAA<br>Lys     | 1104 |
| TCT GAT AAA GAT ATG GAA AAA GAA TAT AGA GAG CAT TTG GCT GTA TAT ATA GAA AAT GAA AAT GAA AAT AAA AAT GAA TTT TTA AAA AA   |   | AC1<br>Thr        |                    | ,                 | T TC<br>a Se       | T TA           | T TT<br>r Ph | C 261          | . wat          | TCC<br>Ser        | TTA<br>Leu     | CCA<br>Pro     | Pro            | GTA<br>Val     | AAA<br>Lys     | ATA<br>Ile   | AAA<br>Lys     | 1152 |
| AAT AAA CAA GCT TAT GAA GAA ATT AGC AAG GAG CAT TTG GCT GTA TAT AGA CAA ATG TYF Glu Glu Ile Ser Lys Glu His Leu Ala Val Tyr 430  AAG ATA CAA ATG TGT AAA AGT GT AAA GCT CCA GGA ATA TGT ATT GAT Lys Ile Gln Met Cys Lys Ser Val Lys Ala Pro Gly Ile Cys Ile Asp 445  GTT GAT AAT GAA GAT TTG TTC TTT ATA GCT GAT AAA AAT AGT TTT TCA 455  GAT GAT TTA TCT AAA AAC GAA AGA ATA GAA TAT AAT ACA CAG AGT AAT ASP ASP Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn 470  TAT ATA GAA AAT GAC TTC CCT ATA AAT GAA TTA ATT TTA GAT ACT GAT TYR Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 495  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT 1536  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT 1536   |   | AAT<br>Asn<br>385 | TTA<br>Leu         | A TT              | A GA<br>u As       | T AA'<br>p Ası |              | * TTC          | TAT            | ACT<br>Thr        | ATA<br>Ile     | Glu            | GAA<br>Glu     | GGG<br>Gly     | TTT<br>Phe     | AAT<br>Asn   | Ile            | 1200 |
| AAG ATA CAA ATG TGT AAA AGT GTT AAA GCT CCA GGA ATA TGT ATT GAT Lys Ile Gln Met Cys Lys Ser Val Lys Ala Pro Gly Ile Cys Ile Asp 445  GTT GAT AAT GAA GAT TTG TTC TTT ATA GCT GAT AAA AAT AGT TTT TCA Val Asp Asn Glu Asp Leu Phe Phe Ile Ala Asp Lys Asn Ser Phe Ser 455  GAT GAT TTA TCT AAA AAC GAA AGA ATA GAA TAT AAT ACA CAG AGT AAT Asp Asp Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn 480  TAT ATA GAA AAT GAC TCC CCT ATA AAT GAA TTA ATT TTA GAT ACT GAT TYr Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 495  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT 1536  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT 1536   |   | TCT<br>Ser        | GAT<br>Asp         | Lys               | A GA'<br>B Asi     |                |              | A AAA<br>1 Lys | GAA<br>Glu     | TAT               | Arg            | GGT<br>Gly     | CAG<br>Gln     | AAT<br>Asn     | AAA<br>Lys     | Ala          | ATA<br>Ile     | 1248 |
| GTT GAT AAT GAA GAT TTG TTC TTT ATA GCT GAT AAA AAT AGT TTT TCA Val Asp Asn Glu Asp Leu Phe Phe Ile Ala Asp Lys Asn Ser Phe Ser 450  GAT GAT TTA TCT AAA AAC GAA AGA ATA GAA TAT AAT ACA CAG AGT AAT Asp Asp Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn 470  TAT ATA GAA AAT GAC TTC CCT ATA AAT GAA TTA ATT TTA GAT ACT GAT TYr Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 485  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asn Thr Glu Ser Leu Thr   |   | TAA<br>Ren        | AAA<br>Lys         | CA/<br>Glr        |                    |                | GAA<br>Glu   | GAA<br>Glu     | ATT            | ser               | AAG<br>Lys     | GAG<br>Glu     | CAT<br>His     | Leu            | Ala            | GTA<br>Val   | TAT<br>Tyr     | 1296 |
| GAT GAT TTA TCT AAA AAC GAA AGA ATA GAA TAT AAT ACA CAG AGT AAT ASP Asp Asp Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn 480  TAT ATA GAA AAT GAC TTC CCT ATA AAT GAA TTA ATT TTA GAT ACT GAT TYr Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 485  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asn Thr Glu Ser Leu Thr  | • | AAG<br>Lys        | ATA<br>Ile         |                   |                    | G TGI<br>Cys   | Lys          | AGT<br>Ser     | val            | AAA<br>Lys        | GCT<br>Ala     | CCA<br>Pro     | Gly            | Ile            | TGT<br>Cys     | ATT lle      | GAT<br>Asp     | 1344 |
| TAT ATA GAA AAT GAC TTC CCT ATA AAT GAA TTA ATT TTA GAT ACT GAT TYR Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 485  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asn Thr Glu Ser Leu Thr   |   | GTT<br>Val        |                    | AAT<br>Asn        | GAA<br>Glu         | GAT<br>Asp     | TTG<br>Leu   | FILE           | TTT<br>Phe     | ATA<br>Ile        | GCT<br>Ala     | Asp            | Lys .          | AAT<br>Asn     | AGT<br>Ser     | TTT<br>Phe   | TCA<br>Ser     | 1392 |
| TAT ATA GAA AAT GAC TTC CCT ATA AAT GAA TTA ATT TTA GAT ACT GAT Tyr Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 485  TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asn Thr Glu Ser Leu Thr   |   |                   | GAT<br>Asp         | TTA<br>Leu        | TCT                | AAA<br>Lys     | 4911         | GAA<br>Glu     | AGA<br>Arg     | ATA<br>Ile        | Glu            | Tyr /          | AAT .<br>Asn " | ACA (          | CAG<br>Gln     | Ser          | Asn            | 1440 |
| TTA ATA AGT AAA ATA GAA TTA CCA AGT GAA AAT ACA GAA TCA CTT ACT Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asn Thr Glu Ser Leu Thr  |   | TAT<br>Tyr        | ATA<br>Ile         | GAA<br>Glu        | AAT<br>Asn         | vab            | TTC<br>Phe   | CCT<br>Pro     | ATA<br>Ile     | AAT<br>Asn        | GLu            | TTA I          | ATT 1          | TTA (<br>Leu i | Asp            | ACT<br>Thr   |                | 1488 |
| 510  |   | TTA<br>Leu.       | ATA<br>Ile         | AGT<br>Ser        | AAA<br>Lys<br>500  | TTE            | GAA<br>Glu   | TTA<br>Leu     | CCA<br>Pro     | AGT<br>Ser<br>505 | GAA /          | AAT 1<br>Asn 1 | ACA (<br>Thr ( | 3lu s          | TCA Ser        | omm.         | ACT<br>Thr     | 1536 |



- 97 -

|   | •                 |                   |                   |                   |                   |                     |                       |                   |                   |                   | _                 | *                 |                       |                   |                   |                       |      |
|---|-------------------|-------------------|-------------------|-------------------|-------------------|---------------------|-----------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-----------------------|-------------------|-------------------|-----------------------|------|
| ٠ | GA'<br>Asj        | r TT:<br>p Phe    | AA:<br>As:<br>51: | n Va.             | A GA:<br>L Asi    | r GTT<br>Val        | r cca                 | Va:<br>520        | l Ty              | r GA/<br>r Glu    | A AAI<br>1 Lys    | A CAI             | A CCC<br>n Pro<br>529 | o Al              | T AT<br>a Il      | A AAA<br>e Lys        | 1584 |
|   | AAJ<br>Lys        | A ATT             | Phe               | r ACA<br>E Thr    | A GAT             | GAA<br>Glu          | A AAT<br>Asn<br>535   | The               | TATO              | TTT<br>Phe        | CAJ<br>Glr        | TAT<br>TY:<br>540 | Let                   | A TA              | C TC<br>r Se      | T CAG<br>r Gln        | 1632 |
|   | ACA<br>Thi        | : Phe             | CCT<br>Pro        | CTA<br>Leu        | GAT<br>Asp        | TATA<br>TILE<br>550 | : Arg                 | GAT<br>Asp        | T ATA             | AGI<br>Ser        | Leu<br>555        | Thr               | TC1                   | TC:               | A TT              | T GAT<br>E Asp<br>560 | 1680 |
|   | GAT<br>Asp        | GCA<br>Ala        | Leu               | A TTA<br>Leu      | TTT<br>Phe<br>565 | Ser                 | AAC<br>Asn            | AAA<br>Lys        | GT1<br>Val        | TAT<br>Tyr<br>570 | Ser               | TTI<br>Phe        | TT1                   | TC!               | F ATO             | G GAT<br>Asp          | 1728 |
|   | TAT<br>Tyr        | ATT<br>Ile        | AAA<br>Lys        | ACT<br>Thr<br>580 | Ala               | AAT<br>Asn          | AAA<br>Lys            | GTG<br>Val        | GTA<br>Val<br>585 | Glu               | GCA<br>Ala        | GGA<br>Gly        | TTA<br>Leu            | TTI<br>Phe<br>590 | Ala               | A GGT<br>A Gly        | 1776 |
|   | TGG<br>Trp        | GTG<br>Val        | AAA<br>Lys<br>595 | Gln               | ATA<br>Ile        | GTA<br>Val          | AAT<br>Asn            | GAT<br>Asp<br>600 | Phe               | GTA<br>Val        | ATC<br>Ile        | GAA<br>Glu        | GCT<br>Ala<br>605     | Asn               | Lys               | AGC<br>Ser            | 1824 |
|   | AAT<br>Asn        | ACT<br>Thr<br>610 | Met               | GAT<br>Asp        | AAA<br>Lys        | ATT<br>Ile          | GCA<br>Ala<br>615     | GAT<br>Asp        | ATA<br>Ile        | TCT<br>Ser        | CTA<br>Leu        | ATT<br>Ile<br>620 | GTT<br>Val            | CCI               | TAT<br>Tyr        | ATA<br>Ile            | 1872 |
|   | GGA<br>Gly<br>625 | Leu               | GCT               | TTA<br>Leu        | AAT<br>Asn        | GTA<br>Val<br>630   | GGA<br>Gly            | AAT<br>Asn        | GAA<br>Glu        | ACA<br>Thr        | GCT<br>Ala<br>635 | AAA<br>Lys        | GGA<br>Gly            | AAT<br>Asn        | TTI               | GAA<br>Glu<br>640     | 1920 |
|   | Asn               | Ala               | Phe               | Glu               | 11e<br>645        | Ala                 | Gly                   | Ala               | Ser               | Ile.<br>650       | Leu               | Leu               | Glu                   | Phe               | Ile<br>655        |                       | 1968 |
|   | GAA<br>Glu        | CTT<br>Leu        | Leu               | ATA<br>Ile<br>660 | Pro               | Val                 | GTT<br>Val            | Gly               | Ala               | Phe               | TTA<br>Leu        | TTA<br>Leu        | GAA<br>Glu            | TCA<br>Ser<br>670 | TAT               | ATT<br>Ile            | 2016 |
|   | GAC<br>Asp        | AAT<br>Asn        | AAA<br>Lys<br>675 | AAT<br>Asn        | AAA<br>Lys        | ATT<br>Ile          | ATT<br>Ile            | AAA<br>Lys<br>680 | ACA<br>Thr        | ATA<br>Ile        | GAT<br>Asp        | AAT<br>Asn        | GCT<br>Ala<br>685     | TTA<br>Leu        | ACT<br>Thr        | AAA<br>Lys            | 2064 |
|   | AGA<br>Arg        | AAT<br>Asn<br>690 | GAA<br>Glu        | AAA<br>Lys        | TGG<br>Trp        | AGT<br>Ser          | GAT<br>Asp<br>695     | ATG<br>Met        | TAC<br>Tyr        | GGA<br>Gly        | TTA<br>Leu        | ATA<br>Ile<br>700 | GTA<br>Val            | GCG<br>Ala        | CAA<br>Gln        | TGG<br>Trp            | 2112 |
|   | CTC<br>Leu<br>705 | TCA<br>Ser        | ACA<br>Thr        | GTT<br>Val        | AAT<br>Asn        | ACT<br>Thr<br>710   | CAA<br>Gln            | TTT<br>Phe        | TAT<br>Tyr        | ACA<br>Thr        | ATA<br>Ile<br>715 | aaa<br>Lys        | GAG<br>Glu            | GGA<br>Gly        | ATG<br>Met        | TAT<br>Tyr<br>720     | 2160 |
|   | AAG<br>Lys        | GCT<br>Ala        | TTA<br>Leu        | AAT<br>Asn        | TAT<br>Tyr<br>725 | CAA<br>Gln          | GCA<br>Ala            | CAA<br>Gln        | Ala               | TTG<br>Leu<br>730 | GAA<br>Glu        | GAA<br>Glu        | ATA<br>Ile            | ATA<br>Ile        | AAA<br>Lys<br>735 | TAC<br>Tyr            | 2208 |
|   | AGA<br>Arg        | TAT<br>Tyr        | AAT<br>Asn        | ATA<br>Ile<br>740 | TAT<br>Tyr        | TCT<br>Ser          | GAA<br>Glu            | Lys               | GAA<br>Glu<br>745 | AAG<br>Lys        | TCA<br>Ser        | AAT<br>Asn        | Ile                   | AAC<br>Asn<br>750 | ATC<br>Ile        | GAT<br>Asp            | 2256 |
|   | TTT<br>Phe        | Asn               | GAT<br>Asp<br>755 | ATA<br>Ile        | AAT<br>Asn        | TCT .<br>Ser :      | Lys                   | CTT<br>Leu<br>760 | AAT<br>Asn        | GAG<br>Glu        | GGT<br>Gly        | Ile               | AAC<br>Asn<br>765     | CAA<br>Gln        | GCT<br>Ala        | ATA<br>Ile            | 2304 |
|   | Asp               | AAT<br>Asn<br>770 | ATA<br>Ile        | AAT .<br>Asn .    | AAT<br>Asn        | Phe :               | ATA A<br>Ile A<br>775 | AAT (<br>Asn (    | GGA<br>Gly        | TGT<br>Cys        | Ser               | GTA<br>Val<br>780 | TCA<br>Ser            | TAT<br>Tyr        | TTA<br>Leu        | ATG<br>Met            | 2352 |



|            |                   |                   |                   |                   |            |                   |                   |                   |                   |            | -          |                   |                   |                   |                   | •    |
|------------|-------------------|-------------------|-------------------|-------------------|------------|-------------------|-------------------|-------------------|-------------------|------------|------------|-------------------|-------------------|-------------------|-------------------|------|
| 785        | <b></b>           | 1100              | 116               | PIO               | 790        | MIA               | vai               | GIU               | Lys               | Leu<br>795 | Leu        | Asp               | Phe               | Asp<br>           | AAT<br>Asn<br>800 | 2400 |
| ACT        | CTC<br>Leu        | AAA<br>Lys        | AAA<br>Lys        | AAT<br>Asn<br>805 | TTG<br>Leu | TTA<br>Leu        | AAT<br>Asn        | TAT<br>Tyr        | ATA<br>Ile<br>810 | GAT<br>Asp | GAA<br>Glu | AAT<br>Asn        | AAA<br>Lys        | TTA<br>Leu<br>815 | TAT<br>Tyr        | 2448 |
| TTG<br>Leu | ATT               | GGA<br>Gly        | AGT<br>Ser<br>820 | GCA<br>Ala        | GAA<br>Glu | TAT<br>Tyr        | GAA<br>Glu        | AAA<br>Lys<br>825 | TCA<br>Ser        | AAA<br>Lys | GTA<br>Val | AAT<br>Asn        | AAA<br>Lys<br>830 | TAC<br>Tyr        | TTG<br>Leu        | 2496 |
| AAA<br>Lys | ACC<br>Thr        | ATT<br>Ile<br>835 | ATG<br>Met        | CCG<br>Pro        | TTT<br>Phe | GAT<br>Asp        | CTT<br>Leu<br>840 | TCA<br>Ser        | ATA<br>Ile        | TAT<br>Tyr | ACC<br>Thr | AAT<br>Asn<br>845 | GAT<br>Asp        | ACA<br>Thr        | ATA<br>Ile        | 2544 |
| CTA<br>Leu | ATA<br>Ile<br>850 | GAA<br>Glu        | ATG<br>Met        | TTT<br>Phe        | AAT<br>Asn | AAA<br>Lys<br>855 | TAT<br>Tyr        | AAT<br>Asn        | AGC<br>Ser        |            |            |                   |                   |                   | .*                | 2574 |

- 98 -

## (2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 858 amino acids (B) TYPE: amino acid

  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Met Pro Val Thr Ile Asn Asn Phe Asn Tyr Asn Asp Pro Ile Asp Asn . 10

Asn Asn Ile Ile Met Met Glu Pro Pro Phe Ala Arg Gly Thr Gly Arg

Tyr Tyr Lys Ala Phe Lys Ile Thr Asp Arg Ile Trp Ile Ile Pro Glu

Arg Tyr Thr Phe Gly Tyr Lys Pro Glu Asp Phe Asn Lys Ser Ser Gly 50 60

Ile Phe Asn Arg Asp Val Cys Glu Tyr Tyr Asp Pro Asp Tyr Leu Asn 65 70 75

Thr Asn Asp Lys Lys Asn Ile Phe Leu Gln Thr Met Ile Lys Leu Phe

Asn Arg Ile Lys Ser Lys Pro Leu Gly Glu Lys Leu Leu Glu Met Ile

Ile Asn Gly Ile Pro Tyr Leu Gly Asp Arg Arg Val Pro Leu Glu Glu

Phe Asn Thr Asn Ile Ala Ser Val Thr Val Asn Lys Leu Ile Ser Asn

Pro Gly Glu Val Glu Arg Lys Lys Gly Ile Phe Ala Asn Leu Ile Ile

Phe Gly Pro Gly Pro Val Leu Asn Glu Asn Glu Thr Ile Asp Ile Gly

Ile Gln Asn His Phe Ala Ser Arg Glu Gly Phe Gly Gly Ile Met Gln



- 99 -

540

Met Lys Phe Cys Pro Glu Tyr Val Ser Val Phe Asn Asn Val Gln Glu 205 Asn Lys Gly Ala Ser Ile Phe Asn Arg Arg Gly Tyr Phe Ser Asp Pro 215 Ala Leu Ile Leu Met His Glu Leu Ile His Val Leu His Gly Leu Tyr Gly Ile Lys Val Asp Asp Leu Pro Ile Val Pro Asn Glu Lys Lys Phe 250 Phe Met Gln Ser Thr Asp Ala Ile Gln Ala Glu Glu Leu Tyr Thr Phe Gly Gly Gln Asp Pro Ser Ile Ile Thr Pro Ser Thr Asp Lys Ser Ile 280 Tyr Asp Lys Val Leu Gln Asn Phe Arg Gly Ile Val Asp Arg Leu Asn Lys Val Leu Val Cys Ile Ser Asp Pro Asn Ile Asn Ile Tyr Lys Asn Lys Phe Lys Asp Lys Tyr Lys Phe Val Glu Asp Ser Glu Gly Lys Tyr Ser Ile Asp Val Glu Ser Phe Asp Lys Leu Tyr Lys Ser Leu Met Phe Gly Phe Thr Glu Thr Asn Ile Ala Glu Asn Tyr Lys Ile Lys 360 Thr Arg Ala Ser Tyr Phe Ser Asp Ser Leu Pro Pro Val Lys Ile Lys 375 Asn Leu Leu Asp Asn Glu Ile Tyr Thr Ile Glu Glu Gly Phe Asn Ile 395 Ser Asp Lys Asp Met Glu Lys Glu Tyr Arg Gly Gln Asn Lys Ala Ile Asn Lys Gln Ala Tyr Glu Glu Ile Ser Lys Glu His Leu Ala Val Tyr 420 Lys Ile Gln Met Cys Lys Ser Val Lys Ala Pro Gly Ile Cys Ile Asp Val Asp Asn Glu Asp Leu Phe Phe Ile Ala Asp Lys Asn Ser Phe Ser Asp Asp Leu Ser Lys Asn Glu Arg Ile Glu Tyr Asn Thr Gln Ser Asn Tyr Ile Glu Asn Asp Phe Pro Ile Asn Glu Leu Ile Leu Asp Thr Asp 485 Leu Ile Ser Lys Ile Glu Leu Pro Ser Glu Asn Thr Glu Ser Leu Thr Asp Phe Asn Val Asp Val Pro Val Tyr Glu Lys Gln Pro Ala Ile Lys Lys Ile Phe Thr Asp Glu Asn Thr Ile Phe Gln Tyr Leu Tyr Ser Gln



#### - 100 -

Thr Phe Pro Leu Asp Ile Arg Asp Ile Ser Leu Thr Ser Ser Phe Asp Asp Ala Leu Leu Phe Ser Asn Lys Val Tyr Ser Phe Phe Ser Met Asp Tyr Ile Lys Thr Ala Asn Lys Val Val Glu Ala Gly Leu Phe Ala Gly Trp Val Lys Gln Ile Val Asn Asp Phe Val Ile Glu Ala Asn Lys Ser Asn Thr Met Asp Lys Ile Ala Asp Ile Ser Leu Ile Val Pro Tyr Ile Gly Leu Ala Leu Asn Val Gly Asn Glu Thr Ala Lys Gly Asn Phe Glu 630 Asn Ala Phe Glu Ile Ala Gly Ala Ser Ile Leu Leu Glu Phe Ile Pro Glu Leu Leu Ile Pro Val Val Gly Ala Phe Leu Leu Glu Ser Tyr Ile Asp Asn Lys Asn Lys Ile Ile Lys Thr Ile Asp Asn Ala Leu Thr Lys Arg Asn Glu Lys Trp Ser Asp Met Tyr Gly Leu Ile Val Ala Gln Trp Leu Ser Thr Val Asn Thr Gln Phe Tyr Thr Ile Lys Glu Gly Met Tyr Lys Ala Leu Asn Tyr Gln Ala Gln Ala Leu Glu Glu Ile Ile Lys Tyr Arg Tyr Asn Ile Tyr Ser Glu Lys Glu Lys Ser Asn Ile Asn Ile Asp Phe Asn Asp Ile Asn Ser Lys Leu Asn Glu Gly Ile Asn Gln Ala Ile Asp Asn Ile Asn Asn Phe Ile Asn Gly Cys Ser Val Ser Tyr Leu Met Lys Lys Met Ile Pro Leu Ala Val Glu Lys Leu Leu Asp Phe Asp Asn Thr Leu Lys Lys Asn Leu Leu Asn Tyr Ile Asp Glu Asn Lys Leu Tyr Leu Ile Gly Ser Ala Glu Tyr Glu Lys Ser Lys Val Asn Lys Tyr Leu 825 Lys Thr Ile Met Pro Phe Asp Leu Ser Ile Tyr Thr Asn Asp Thr Ile Leu Ile Glu Met Phe Asn Lys Tyr Asn Ser

#### (2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1644 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

- 101 -

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..1644

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 23:

|                   | (x.                | i) S              | EQUE              | NCE 1             | DESC                | RIPTI                 | ON:               | SEQ                | ID 1              | 10: 2             | 23:               |                   |                       |                   |                     |       |
|-------------------|--------------------|-------------------|-------------------|-------------------|---------------------|-----------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-----------------------|-------------------|---------------------|-------|
| Me                | G CC<br>t Pro<br>1 | A GT<br>o Va      | T AC              | A ATA             | A AAT<br>e Asi<br>5 | r aal<br>1 Asr        | TT:               | T AA?<br>e Asr     | TAT<br>TY1        | Asr               | GA?               | CCT<br>Pro        | r ar                  | T GA<br>B Asp     | T AAT<br>p Asn<br>5 | 48    |
| AA:<br>Asi        | r AA:<br>n Asi     | r AT              | T AT              | e met             | G ATO               | GAG<br>Glu            | CC:               | r cca<br>Pro<br>25 | ) Phe             | GCG<br>Ala        | AGA<br>Arg        | A GGT<br>g Gly    | 7 ACC<br>7 Th:<br>3 C | c Gly             | G AGA<br>Y Arg      | 96    |
| тух               | r Tyr              | 35<br>Lys         | S Ala             | a Pne             | : Lys               | ille                  | Thr<br>40         | Asp                | Arg               | Ile               | Trp               | 11e               | : Il€                 | Pro               | G GAA               | 144   |
| Arc               | 50                 | ini               | . Phe             | e GIY             | Tyr                 | 55<br>Lys             | Pro               | Glu                | Asp               | Phe               | Asn<br>60         | Lys               | Ser                   | Ser               | GGT                 | 192   |
| ATT<br>Ile<br>65  | Pne                | AAI<br>Asn        | AGA<br>Arg        | GAT<br>Asp        | GTT<br>Val<br>70    | Cys                   | GAA<br>Glu        | TAT<br>Tyr         | TAT               | GAT<br>Asp<br>75  | CCA<br>Pro        | GAT<br>Asp        | TAC<br>Tyr            | TTA<br>Leu        | AAT<br>Asn<br>80    | . 240 |
| ACT<br>Thr        | ' AAT<br>' Asn     | GAT<br>Asp        | 'AAA<br>Lys       | AAG<br>Lys<br>85  | Asn                 | ATA<br>Ile            | TTT               | TTA<br>Leu         | CAA<br>Gln<br>90  | ACA<br>Thr        | ATG<br>Met        | ATC<br>Ile        | AAG<br>Lys            | TTA<br>Leu<br>95  | TTT                 | 288   |
| Asn               | Arg                | 11e               | 100               | Ser               | Lys                 | Pro                   | Leu               | Gly<br>105         | Glu               | Lys               | Leu               | Leu               | Glu<br>110            | Met               | ATT<br>Ile          | 336   |
| ATA<br>Ile        | AAT<br>Asn         | GGT<br>Gly<br>115 | ATA<br>Ile        | · CCT<br>Pro      | TAT<br>Tyr          | CTT<br>Leu            | GGA<br>Gly<br>120 | GAT<br>Asp         | AGA<br>Arg        | CGT<br>Arg        | GTT<br>Val        | CCA<br>Pro<br>125 | CTC<br>Leu            | GAA<br>Glu        | GAG<br>Glu          | 384   |
| TTT<br>Phe        | AAC<br>Asn<br>130  | ACA<br>Thr        | AAC<br>Asn        | ATT               | GCT<br>Ala          | AGT<br>Ser<br>135     | GTA<br>Val        | ACT<br>Thr         | GTT<br>Val        | AAT<br>Asn        | AAA<br>Lys<br>140 | TTA<br>Leu        | ATC<br>Ile            | AGT<br>Ser        | AAT<br>Asn          | 432   |
| CCA<br>Pro<br>145 | GGA<br>Gly         | GAA<br>Glu        | GTG<br>Val        | GAG<br>Glu        | CGA<br>Arg<br>150   | AAA<br>Lys            | AAA<br>Lys        | GGT<br>Gly         | ATT<br>Ile        | TTC<br>Phe<br>155 | GCA<br>Ala        | AAT<br>Asn        | TTA<br>Leu            | ATA<br>Ile        | ATA<br>Ile<br>160   | 480   |
| TTT<br>Phe        | GGA<br>Gly         | CCT<br>Pro        | GGG<br>Gly        | CCA<br>Pro<br>165 | GTT<br>Val          | TTA<br>Leu            | AAT<br>Asn        | GAA<br>Glu         | AAT<br>Asn<br>170 | GAG<br>Glu        | ACT<br>Thr        | ATA<br>Ile        | GAT<br>Asp            | ATA<br>Ile<br>175 | GGT<br>Gly          | 528   |
| ATA<br>Ile        | CAA<br>Gln         | AAT<br>Asn        | CAT<br>His<br>180 | TTT<br>Phe        | GCA<br>Ala          | TCA<br>Ser            | AGG<br>Arg        | GAA<br>Glu<br>185  | GGC<br>Gly        | TTC<br>Phe        | GGG<br>Gly        | GGT<br>Gly        | ATA<br>Ile<br>190     | ATG<br>Met        | CAA<br>Gln          | 576   |
| ATG<br>Met        | AAG<br>Lys         | TTT<br>Phe<br>195 | TGC<br>Cys        | CCA<br>Pro        | GAA<br>Glu          | Tyr                   | GTA<br>Val<br>200 | AGC<br>Ser         | GTA<br>Val        | TTT<br>Phe        | AAT<br>Asn        | AAT<br>Asn<br>205 | GTT<br>Val            | CAA<br>Gln        | GAA<br>Glu          | 624   |
| Asn               | AAA<br>Lys<br>210  | GGC<br>Gly        | GCA<br>Ala        | AGT<br>Ser        | Ile                 | TTT .<br>Phe .<br>215 | AAT<br>Asn        | AGA<br>Arg         | CGT<br>Arg        | Gly               | TAT<br>Tyr<br>220 | TTT<br>Phe        | TCA<br>Ser            | GAT<br>Asp        | CCA<br>Pro          | 672   |



# - 102 -

|                   |            |                   |                   |                   |                   |             |                   |                   | ٠.                |                   |            |                   | •                 |                   |                   |                                       |      |   |
|-------------------|------------|-------------------|-------------------|-------------------|-------------------|-------------|-------------------|-------------------|-------------------|-------------------|------------|-------------------|-------------------|-------------------|-------------------|---------------------------------------|------|---|
| GCC<br>Ala<br>225 | TTG<br>Leu | ATA<br>Ile        | TTA<br>Leu        | ATG<br>Met        | CAT<br>His<br>230 | GAA<br>Glu  | CTT<br>Leu        | ATA<br>Ile        | CAT               | GTT<br>Val<br>235 | Leu        | CAT               | GGA<br>Gly        | TTA<br>Leù        | TAT<br>Tyr<br>240 |                                       | 720  |   |
| GGC<br>Gly        | ATT<br>Ile | AAA<br>Lys        | GTA<br>Val        | GAT<br>Asp<br>245 | GAT<br>Asp        | TTA<br>Leu  | CCA<br>Pro        | ATT<br>Ile        | GTA<br>Val<br>250 | CCA<br>Pro        | AAT<br>Asn | GAA<br>Glu        | AAA<br>Lys        | AAA<br>Lys<br>255 | TTT<br>Phe        |                                       | 768  |   |
| TTT<br>Phe        | ATG<br>Met | CAA<br>Gln        | TCT<br>Ser<br>260 | ACA<br>Thr        | GAT<br>Asp        | GCT.<br>Ala | ATA<br>Ile        | CAG<br>Gln<br>265 | GCA<br>Ala        | GAA<br>Glu        | GAA<br>Glu | CTA<br>Leu        | TAT<br>Tyr<br>270 | ACA<br>Thr        | TTT<br>Phe        | •                                     | 816  |   |
| GGA<br>Gly        | GGA<br>Gly | CAA<br>Gln<br>275 | GAT<br>Asp        | Pro               | AGC<br>Ser        | ATC<br>Ile  | ATA<br>Ile<br>280 | Thr               | CCT               | TCT<br>Ser        | ACG<br>Thr | GAT<br>Asp<br>285 | AAA<br>Lys        | AGT<br>Ser        | ATC Ile           |                                       | 864  |   |
|                   |            |                   |                   | TTG<br>Leu        |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 912  |   |
| AAG<br>Lys<br>305 | GTT<br>Val | TTA<br>Leu        | GTT<br>Val        | TGC<br>Cys        | ATA<br>Ile<br>310 | TCA<br>Ser  | GAT<br>Asp        | CCT<br>Pro        | AAC<br>Asn        | ATT<br>Ile<br>315 | AAT<br>Asn | ATT<br>Ile        | AAT<br>Asn        | ATA<br>Ile        | TAT<br>Tyr<br>320 |                                       | 960  |   |
|                   |            |                   |                   | AAA<br>Lys<br>325 | Asp               |             |                   |                   |                   | Val               |            |                   |                   |                   |                   |                                       | 1008 |   |
|                   |            |                   |                   | GAT<br>Asp        |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1056 |   |
|                   |            |                   |                   | ACA<br>Thr        |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   | · · · · · · · · · · · · · · · · · · · | 1104 |   |
|                   |            | Ala               |                   | TAT<br>Tyr        |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1152 | , |
|                   |            |                   |                   | AAT<br>Asn        |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1200 |   |
|                   |            |                   |                   | ATG<br>Met<br>405 |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1248 | - |
|                   |            |                   |                   | TAT<br>Tyr        |                   |             |                   |                   |                   |                   |            |                   |                   | Val               |                   |                                       | 1296 |   |
|                   |            |                   |                   | TGT<br>Cys        |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1344 |   |
|                   |            |                   |                   | GAT<br>Asp        | Leu               |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1392 |   |
|                   |            |                   |                   | AAA<br>Lys        |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1440 |   |
|                   |            |                   |                   | GAC<br>Asp<br>485 |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       | 1488 |   |
|                   |            |                   |                   |                   |                   |             |                   |                   |                   |                   |            |                   |                   |                   |                   |                                       |      |   |



- 103 -

| TTA<br>Leu | ATA<br>Ile        | AGT<br>Ser        | AAA<br>Lys<br>500 | ATA<br>Ile | GAA<br>Glu | TTA<br>Leu        | CCA<br>Pro        | AGT<br>Ser<br>505 | GAA<br>Glu | AAT<br>Asn | ACA<br>Thr        | GAA<br>Glu        | TCA<br>Ser<br>510 | CTT<br>Leu | ACT<br>Thr | 1536 |
|------------|-------------------|-------------------|-------------------|------------|------------|-------------------|-------------------|-------------------|------------|------------|-------------------|-------------------|-------------------|------------|------------|------|
| GAT<br>Asp | TTT<br>Phe        | AAT<br>Asn<br>515 | GTA<br>Val        | GAT<br>Asp | GTT<br>Val | CCA<br>Pro        | GTA<br>Val<br>520 | TAT<br>Tyr        | GAA<br>Glu | AAA<br>Lys | CAA<br>Gln        | CCC<br>Pro<br>525 | GCT<br>Ala        | ATA<br>Ile | AAA<br>Lys | 1584 |
| AAA<br>Lys | ATT<br>Ile<br>530 | TTT<br>Phe        | ACA<br>Thr        | GAT<br>Asp | GAA<br>Glu | TAA<br>Asn<br>535 | ACC<br>Thr        | ATC<br>Ile        | TTT<br>Phe | CAA<br>Gln | TAT<br>Tyr<br>540 | TTA<br>Leu        | TAC<br>Tyr        | TCT<br>Ser | CAG<br>Gln | 1632 |
|            | TTT<br>Phe        |                   |                   |            |            |                   |                   |                   |            |            |                   |                   |                   |            |            | 1644 |

#### (2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 548 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Met Pro Val Thr Ile Asn Asn Phe Asn Tyr Asn Asp Pro Ile Asp Asn Asn Asn Ile Ile Met Met Glu Pro Pro Phe Ala Arg Gly Thr Gly Arg 20 Tyr Tyr Lys Ala Phe Lys Ile Thr Asn Asp Phe Asn Lys Ser Ser Gly 55 Phe Asn Asp Phe Asn Lys Ser Ser Gly Fro Glu Tyr Tyr Tyr Asp Pro Asp Tyr Leu Asn 80 Thr Asn Asp Lys Lys Asn Ile Phe Leu Gln Thr Met Ile Lys Leu Phe

85 90 95

Asn Arg Ile Lys Ser Lys Pro Leu Gly Glu Lys Leu Leu Glu Met Ile 100 105 110

Ile Asn Gly Ile Pro Tyr Leu Gly Asp Arg Arg Val Pro Leu Glu Glu 115 125

Phe Asn Thr Asn Ile Ala Ser Val Thr Val Asn Lys Leu Ile Ser Asn 130 135 140

Pro Gly Glu Val Glu Arg Lys Lys Gly Ile Phe Ala Asn Leu Ile Ile 145 150 155 160

Phe Gly Pro Gly Pro Val Leu Asn Glu Asn Glu Thr Ile Asp Ile Gly 165 170 175

Ile Gln Asn His Phe Ala Ser Arg Glu Gly Phe Gly Gly Ile Met Gln 180 185 190

Met Lys Phe Cys Pro Glu Tyr Val Ser Val Phe Asn Asn Val Gln Glu 195 200 205



# - 104 -

|    | •          |            |            |            |              |              |             |            |            | •          |            |       |            |              |            |            |  |
|----|------------|------------|------------|------------|--------------|--------------|-------------|------------|------------|------------|------------|-------|------------|--------------|------------|------------|--|
|    | •          |            |            |            |              |              | 213         | )          |            | ,          |            | 220   |            |              | ł          | Pro        |  |
|    | 22.        | •          |            | •          |              | His<br>230   |             |            | •          |            | 235        | ,     |            |              |            | 240        |  |
|    | Gly        | / Ile      | Lys        | Val        | . Asp<br>245 | Asp          | Leu         | Pro        | Ile        | Val<br>250 | Pro        | Asn   | Glu        | Lys          | Lys<br>255 | Phe        |  |
|    | Phe        | e Met      | Gln        | Ser<br>260 | Thr          | Asp          | Ala         | Ile        | Gln<br>265 | Ala        | Glu        | Glu   | Leu        | Tyr<br>270   | Thr        | Phe        |  |
|    | Gly        | Gly        | Gln<br>275 | Asp        | Pro          | Ser          | Ile         | Ile<br>280 | Thr        | Pro        | Ser        | Thr   | Asp<br>285 | Lys          | Ser        | Ile        |  |
|    |            | 290        |            |            |              | Gln          | 295         |            |            |            |            | 300   |            |              |            | •          |  |
|    | Lys<br>305 | Val        | Leu        | Val        | Cys          | Ile<br>310   | Ser         | Asp        | Pro        | Asn        | Ile<br>315 | Asn   | Ile        | Asn          | Ile        | Tyr<br>320 |  |
|    | Lys        | Asn        | Lys        | Phe        | Lys<br>325   | Asp          | Lys         | Tyr        | Lys        | Phe<br>330 | Val        | Glu   | Asp        | Ser          | Glu<br>335 | Gly        |  |
| ٠. |            |            | • .        | 340        |              | Val          |             |            | 345        |            |            |       |            | 350          |            |            |  |
|    | Met        | Phe        | Gly<br>355 | Phe        | Thr          | Glu          | Thr         | Asn<br>360 | Ile        | Ala        | Glu        | Asn   | Tyr<br>365 | Lys          | Ile        | Lys        |  |
|    |            | 370        |            |            | ٠.           | Phe          | 3/5         |            |            |            |            | 380   |            |              | •          | _          |  |
|    | 203        |            |            |            |              | Glu<br>390   |             |            |            |            | 395        |       |            |              |            | 400        |  |
|    | . :        |            | •          |            | 403          | Glu          |             |            |            | 410        |            |       |            |              | 415        |            |  |
| .` |            |            | ,          | 420        |              | Glu          |             |            | 425        |            |            |       |            | 430          |            |            |  |
|    | . •        | ,          | 433        | •          |              | Lys :        |             | 440        |            |            |            |       | 445        |              |            |            |  |
|    |            | 450        |            |            |              |              | 433         |            |            |            |            | 460   |            |              |            |            |  |
|    | Asp<br>465 | Asp        | Leu        | Ser        | Lys .        | Asn (<br>470 | Glu         | Arg        | Ile        | Glų        | Tyr<br>475 | Asn   | Thr        | Gln :        |            | Asn<br>180 |  |
|    |            | •          |            | •          | 485          | Phe I        |             |            |            | 490        |            |       |            | •            | 195        | ,          |  |
|    | Leu        | Ile        | Ser        | Lys<br>500 | Ile (        | Glu I        | Leu :       | Pro        | Ser<br>505 | Glu .      | Asn '      | Thr   |            | Ser 1<br>510 | Leu 7      | Thr        |  |
|    |            |            | 212        | ÷          |              | Val I        | ;           | 520        |            |            |            |       | 525        |              |            | _          |  |
|    | Lys        | Ile<br>530 | Phe        | Thr /      | qa <i>A</i>  | Glu A        | Asn (<br>35 | Thr :      | Ile        | Phe (      | Gln '      | Tyr : | Leu 1      | Tyr S        | Ser (      | Sln        |  |
|    | Thr<br>545 | Phe        | Pro        | Leu        |              |              |             |            |            |            |            |       |            |              |            |            |  |

- 105 -

# (2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:

  (A) LENGTH: 2616 base pairs

  (B) TYPE: nucleic acid

  (C) STRANDEDNESS: double

  (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:

  - (A) NAME/KEY: CDS
    (B) LOCATION:1..2616

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

| ATG<br>Met     | Glr.       | TTO<br>Phe       | C GTC             | AAC<br>L Asn     | AAG<br>Lys | CAG<br>Gln | TTC<br>Phe       | AAC<br>: Aan      | TAT<br>Tyx<br>10 | Lys        | GAC<br>Asp | Pro              | GTA<br>Val        | A AAC<br>Asr<br>15 | GGT<br>Gly |   | 48  |
|----------------|------------|------------------|-------------------|------------------|------------|------------|------------------|-------------------|------------------|------------|------------|------------------|-------------------|--------------------|------------|---|-----|
| GTT<br>Val     | GAC<br>Asp | ATI<br>Ile       | GCC<br>Ala        | ııyı             | ATC<br>Ile | AAA<br>Lys | ATT<br>Ile       | CCA<br>Pro<br>25  | Asn              | GCC        | GGC<br>Gly | CAC<br>Gln       | ATC<br>Met        | Gln                | CCG<br>Pro |   | 96  |
| GTG<br>Val     | AAG<br>Lys | GCT<br>Ala<br>35 | Pne               | Lys              | ATT        | CAT        | AAC<br>Asn<br>40 | Lys               | ATC              | TGG<br>Trp | GTI<br>Val | ATT<br>Ile<br>45 | Pro               | GAA<br>Glu         | CGC        |   | 144 |
| Asp            | 50         | Pne              | Inr               | ASN              | PIO        | 55<br>55   | Glu              | Gly               | qaA              | Leu        | Asn<br>60  | Pro              | Pro               | Pro                | GAA<br>Glu | · | 192 |
| 65             | гуя        | GIN              | vai               | PTO              | 70         | ser        | Tyr              | Tyr               | Asp              | Ser<br>75  | Thr        | Tyr              | Leu               | Ser                | 80         |   | 240 |
| Asp            | ASII       | GIU              | Lys               | GAT<br>Asp<br>85 | Asn        | Tyr        | Leu              | Lys               | Gly<br>90        | Val        | Thr        | Lys              | Leu               | Phe<br>95          | Glu        |   | 288 |
| Arg            | me         | lyr              | 100               | ACT              | Asp        | Leu        | GLY              | Arg<br>105        | Met              | Leu        | Leu        | Thr              | Ser<br>110        | Ile                | Val        |   | 336 |
| Arg            | GIÀ        | 115              | Pro               | TTT<br>Phe       | Trp        | GIA        | Gly<br>120       | Ser               | Thr              | Ile        | Asp        | Thr<br>125       | Glu               | Leu                | Lys        | ; | 384 |
| vai            | 11e<br>130 | Asp              | Thr               | AAC<br>Asn       | Cys        | 11e<br>135 | Asn              | Val               | Ile              | Gln        | Pro<br>140 | Asp              | Gly               | Ser                | Tyr        | • | 432 |
| Arg<br>145     | ser        | GIU              | GIU               | CTT<br>Leu       | Asn<br>150 | Leu        | Val              | Ile               | Ile              | Gly<br>155 | Pro        | Ser              | Ala               | Asp .              | Ile<br>160 |   | 480 |
| TIE (          | GIU        | Pne ·            | GIU               | 165              | Lys        | Ser        | Phe              | Gly               | His<br>170       | Glu        | Val        | Leu              | Asn               | Leu<br>175         |            |   | 528 |
| CGT /<br>Arg / | AAC<br>Asn | GGT<br>Gly       | TAC<br>Tyr<br>180 | GGC<br>Gly       | TCT<br>Ser | ACT<br>Thr | Gln              | TAC<br>Tyr<br>185 | ATT (            | CGT<br>Arg | TTC<br>Phe | Ser              | CCA<br>Pro<br>190 | GAC<br>Asp         | TTC<br>Phe | 5 | 576 |



# - 106 -

|            |                   |                   |                   |                   | GAG<br>Glu        |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 624  |
|------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------------|-------------------|-------------------|-------------------|-------------------|------------|---|------|
|            |                   |                   |                   |                   | GCA<br>Ala        |                   |                   |                   |                   |            |                   |                   |                   |                   |            | - | 672  |
|            |                   |                   |                   |                   | CAT<br>His<br>230 |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 720  |
|            |                   |                   |                   |                   | AAC<br>Asn        |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 768  |
|            |                   |                   |                   |                   | GAA<br>Glu        |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 816  |
|            |                   |                   |                   |                   | CAG<br>Gln        |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 864  |
|            |                   |                   |                   |                   | GCA<br>Ala        |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 912  |
|            |                   |                   |                   |                   | TTA<br>Leu<br>310 |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 960  |
|            |                   |                   |                   |                   | GAT<br>Asp        |                   |                   |                   |                   |            |                   |                   |                   |                   |            | • | 1008 |
|            |                   |                   |                   |                   | TAC<br>Tyr        |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 1056 |
| AAT<br>Asn | TTT               | GTT<br>Val<br>355 | AAG<br>Lys        | TTT<br>Phe        | TTT<br>Phe        | AAA<br>Lys        | GTA<br>Val<br>360 | CTT<br>Leu        | AAC<br>Asn        | AGA<br>Arg | AAA<br>Lys        | ACA<br>Thr<br>365 | TAT<br>Tyr        | TTG<br>Leu        | AAT<br>Asn |   | 1104 |
| TTT<br>Phe | GAT<br>Asp<br>370 | AAA<br>Lys        | GCC<br>Ala        | GTA<br>Val        | TTT               | AAG<br>Lys<br>375 | ATA<br>Ile        | AAT<br>Asn        | ATA<br>Ile        | GTA<br>Val | CCT<br>Pro<br>380 | AAG<br>Lys        | GTA<br>Val        | AAT<br>Asn        | TAC        |   | 1152 |
|            |                   |                   |                   |                   | TTT<br>Phe<br>390 |                   |                   |                   |                   |            |                   |                   |                   |                   |            |   | 1200 |
| TTT<br>Phe | TAA<br>Asn        | GGT<br>Gly        | CAA<br>Gln        | AAT<br>Asn<br>405 | ACA<br>Thr        | GAA<br>Glu        | ATT<br>Ile        | AAT<br>Asn        | AAT<br>Asn<br>410 | ATG<br>Met | AAT<br>Asn        | TTT               | ACT<br>Thr        | AAA<br>Lys<br>415 | CTA<br>Leu |   | 1248 |
| AAA<br>Lys | TAA<br>neA        | TTT<br>Phe        | ACT<br>Thr<br>420 | GGA<br>Gly        | TTG<br>Leu        | TTT<br>Phe        | GAA<br>Glu        | TTT<br>Phe<br>425 | TAT               | AAG<br>Lys | TTG<br>Leu        | CTA<br>Leu        | TGT<br>Cys<br>430 | GTA<br>Val        | AGA<br>Arg |   | 1296 |
| GGG<br>Gly | ATA<br>Ile        | ATA<br>Ile<br>435 | ACT<br>Thr        | TCT<br>Ser        | AAA<br>Lys        | ACT<br>Thr        | AAA<br>Lys<br>440 | TCA<br>Ser        | TTA<br>Leu        | GAT<br>Asp | AAA<br>Lys        | GGA<br>Gly<br>445 | TAC<br>Tyr        | AAT<br>Asn        | AAG<br>Lys |   | 1344 |
| GCA<br>Ala | TTA<br>Leu<br>450 | TAA               | GAT<br>Asp        | TTA<br>Leu        | TGT<br>Cys        | ATC<br>Ile<br>455 | AAA<br>Lys        | GTT<br>Val        | AAT<br>Asn        | AAT<br>Asn | TGG<br>Trp<br>460 | GAC<br>Asp        | TTG<br>Leu        | TTT<br>Phe        | TTT<br>Phe |   | 1392 |

- 107 -

| A0<br>Se<br>46    | T E                | CT T         | CA (         | GAA<br>Glu        | GAT<br>Asp         | C AA<br>D As:<br>47 | n Pn              | T AC                 | T A                | AT GI              | sp L               | TA A<br>eu A<br>75 | \AT<br>ne/     | AA#<br>Lys        | A GG<br>5 Gl      | A G#<br>y Gl       | AA<br>Lu       | GAA<br>Glu<br>480 |     | 1440 |
|-------------------|--------------------|--------------|--------------|-------------------|--------------------|---------------------|-------------------|----------------------|--------------------|--------------------|--------------------|--------------------|----------------|-------------------|-------------------|--------------------|----------------|-------------------|-----|------|
| AT<br>Il          | T AC               | A T          | CT (         | GAT<br>Asp        | ACT<br>Thr<br>485  | Wal                 | T AT              | A GA<br>e Gl         | A GO<br>u Al       | A GO<br>a Al<br>49 | la G               | AA G<br>lu G       | AA<br>lu       | AAT<br>Asn        | AT Il             | T AC<br>e Se<br>49 | r              | TTA<br>Leu        |     | 1488 |
| GA<br>As          | T TI<br>p Le       | 'A A'        | בפ כ         | CAA<br>Sln<br>SOO | CAA<br>Gln         | TA:                 | TA'               | T TT.                | A AC<br>u Th<br>50 | r Ph               | T A                | AT T<br>sn P       | TT<br>he       | GAT<br>Asp        | AA:<br>Asi<br>510 | n Gl               | A              | CCT<br>Pro        |     | 1536 |
| GA<br>Gl          | A AA<br>u As       | T AT<br>n II | LE 3         | CA<br>er          | ATA<br>Ile         | GAZ<br>Glu          | A AA'<br>1 Asi    | r cr<br>n Lei<br>520 | u Se               | A AG               | T GJ               | AC A'<br>Sp I      | TT<br>le       | ATA<br>Ile<br>525 | GG(               | C CA               | A i            | TTA<br>Leu        |     | 1584 |
| GA.<br>Gl         | A CT<br>u Le<br>53 | ט ויוכ       | G C          | TO                | AAT<br>Asn         | ATA                 | GA/<br>Glu<br>535 | A AGI<br>1 Arg       | A TT               | T CC<br>e Pr       | T AF               | n G                | GA<br>ly<br>40 | AAA<br>Lys        | AAC<br>Lys        | TA'                | T (            | GAG<br>Glu        |     | 1632 |
| TT/<br>Let<br>545 | ı AS               | r AA<br>P Ly | A T          | AT<br>yr          | ACT<br>Thr         | ATG<br>Met<br>550   | Pne               | CAT<br>His           | TA:                | r Le               | T CO<br>u Ar<br>55 | g A)               | T<br>la        | CAA<br>Gln        | GAA<br>Glu        | TT                 | <b>e</b> (     | AA<br>Slu<br>660  |     | 1680 |
| 1115              | . G1)              | у пу         | S 51         | er .              | 565                | iie                 | ATS               | TTA<br>Leu           | Thi                | 570                | n Se               | r Va               | al A           | Asn               | Glu               | Ala<br>575         | a I            | eu                |     | 1728 |
| TTA<br>Leu        | AA?<br>Asr         | CC<br>Pr     | U St         | GT<br>er 1<br>80  | CGT<br>Arg         | GTT<br>Val          | TAT               | ACA<br>Thr           | Phe<br>585         | Phe                | T TC<br>Se         | T TC<br>r Se       | A C            | SAC<br>Asp        | TAT<br>Tyr<br>590 | GT#<br>Val         | A<br>L         | AG<br>ys          | ٠   | 1776 |
| nys               | vai                | 59!          | и <b>Б</b> у | /s /              | чта                | Inr                 | GIU               | GCA<br>Ala<br>600    | Ala                | Met                | : Ph               | e Le               | u 6            | 31y<br>305        | Trp               | Val                | G              | lu                |     | 1824 |
| G111              | 610                | va.          | г ту         | T #               | rsb                | Pne                 | 615               | GAT<br>Asp           | Glu                | Thr                | Sei                | 62                 | u V<br>0       | al                | Ser               | Thr                | T.             | hr                | :   | 1872 |
| 625               | Dys                | 116          | H            | a P               | sp                 | 630                 | THE               | ATA<br>Ile           | He                 | Ile                | Pro<br>635         | <b>Ty</b> :        | r I            | le (              | Gly               | Pro                | A.             | la<br>10          | . 1 | L920 |
| Deu               | ABII               | 116          | . GI         | y A<br>6          | 45                 | met                 | reu               | TAT<br>Tyr           | Lys                | Asp<br>650         | Asp                | Phe                | e V            | al (              | Gly               | Ala<br>655         | Le             | eu.               | 1   | .968 |
| 116               | PILE               | Ser          | 66           | <b>y A</b>        | Ta /               | Val                 | IIE               | CTG<br>Leu           | Leu<br>665         | Glu                | Phe                | Ile                | <b>?</b> P:    | ro (              | 31u<br>570        | Ile                | AJ             | .a                | 2   | 016  |
| 116               | PIO                | 675          | net          | u G               | TA 1               | inr                 | Pne               | GCA<br>Ala<br>680    | Leu                | Val                | Ser                | Тут                | 6 i            | le A<br>85        | lla .             | Asn                | Ly             | 'S                | 2   | 064  |
| VAI               | CTA<br>Leu<br>690  | ACC<br>Thr   | GT:<br>Val   | r c               | AA A<br>ln 1       | inr .               | ATA<br>Ile<br>695 | GAT<br>Asp           | AAT<br>Asn         | GCT<br>Ala         | TTA<br>Leu         | AGT<br>Ser<br>700  | L              | AA A<br>A ay      | GA .<br>.rg .     | AAT<br>Asn         | GA<br>G1       | A<br>u            | 2   | 112  |
| AAA<br>Lys<br>705 | TGG<br>Trp         | GAT<br>Asp   | GAC<br>Glu   | 5 G:              | ar i               | Yr I                | AAA<br>Lys        | TAT<br>Tyr           | ATA<br>Ile         | GTA<br>Val         | ACA<br>Thr<br>715  | AAT<br>Asn         | T              | G T<br>Tp L       | TA (              | GCA<br>Ala         | AA<br>Ly<br>72 | s                 | 2   | 160  |
| GTT<br>Val        | TAA<br>Ran         | ACA<br>Thr   | CAG<br>Gln   | ΓTÝ               | TT G<br>le A<br>25 | AT (                | CTA .<br>Leu      | ATA .<br>Ile .       | AGA<br>Arg         | AAA<br>Lys<br>730  | AAA<br>Lys         | ATG<br>Met         | A.P.<br>Ly     | LA G<br>'s G      | lu A              | GCT<br>Ala<br>735  | TT<br>Le       | A<br>u            | 22  | 208  |



#### - 108 -

| 010               |                   |                   | 740               | GIU               | . WIG             | . Inc             | гÀа               | 745               | i ite             | lle               | Asn               | Tyr               | Gln<br>750        | Tyr               | AAT<br>Asn        | 2256 |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|
| CAA<br>Gln        | TAT               | ACT<br>Thr<br>755 | GAG<br>Glu        | GAA<br>Glu        | GAG<br>Glu        | AAA<br>Lys        | AAT<br>Asn<br>760 | Asn               | ATT               | AAT<br>Asn        | TTT<br>Phe        | AAT<br>Asn<br>765 | ATT               | GAT<br>Asp        | GAT<br>Asp        | 2304 |
| TTA<br>Leu        | AGT<br>Ser<br>770 | TCG<br>Ser        | AAA<br>Lys        | CTT<br>Leu        | AAT<br>Asn        | GAG<br>Glu<br>775 | TCT<br>Ser        | ATA<br>Ile        | AAT<br>Asn        | AAA<br>Lys        | GCT<br>Ala<br>780 | Met               | ATT<br>Ile        | AAT<br>Asn        | ATA<br>Ile        | 2352 |
| AAT<br>Asn<br>785 | AAA<br>Lys        | TTT               | TTG               | AAT<br>Asn        | CAA<br>Gln<br>790 | TGC<br>Cys        | TCT<br>Ser        | GTT<br>Val        | TCA<br>Ser        | TAT<br>Tyr<br>795 | TTA<br>Leu        | ATG<br>Met        | AAT<br>Asn        | TCT               | ATG<br>Met<br>800 | 2400 |
| ATC<br>Ile        | CCT<br>Pro        | TAT               | GGT<br>Gly        | GTT<br>Val<br>805 | AAA<br>Lys        | CGG<br>Arg        | TTA<br>Leu        | GAA<br>Glu        | GAT<br>Asp<br>810 | Phe               | GAT<br>Asp        | GCT<br>Ala        | AGT<br>Ser        | CTT<br>Leu<br>815 | AAA<br>Lys        | 2448 |
| GAT<br>Asp        | GCA<br>Ala        | TTA<br>Leu        | TTA<br>Leu<br>820 | AAG<br>Lys        | TAT<br>Tyr        | ATA<br>Ile        | TAT<br>Tyr        | GAT<br>Asp<br>825 | AÀT<br>Asn        | AGA<br>Arg        | GGA<br>Gly        | ACT<br>Thr        | TTA<br>Leu<br>830 | ATT<br>Ile        | GGT<br>Gly        | 2496 |
| CAA<br>Gln        | GTA<br>Val        | GAT<br>Asp<br>835 | AGA<br>Arg        | TTA<br>Leu        | AAA<br>Lys        | GAT<br>Asp        | AAA<br>Lys<br>840 | GTT<br>Val        | AAT<br>Asn        | AAT<br>Asn        | ACA<br>Thr        | CTT<br>Leu<br>845 | AGT<br>Ser        | ACA<br>Thr        | GAT<br>Asp        | 2544 |
| 446               | CCT<br>Pro<br>850 | TTT<br>Phe        | CAG<br>Gln        | CTT<br>Leu        | TCC<br>Ser        | AAA<br>Lys<br>855 | TAC<br>Tyr        | GTA<br>Val        | GAT<br>Asp        | AAT<br>Asn        | CAA<br>Gln<br>860 | AGA<br>Arg        | TTA<br>Leu        | TTA<br>Leu        | TCT<br>Ser        | 2592 |
| ACA<br>Thr<br>865 | TTT<br>Phe        | ACT<br>Thr        | GAA<br>Glu        | TAT<br>Tyr        | ATT<br>Ile<br>870 | AAG<br>Lys        | TAA<br>*          | ,                 |                   |                   |                   |                   |                   |                   | ·                 | 2616 |
|                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |      |

## (2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 872 amino acids
  - (B) TYPE: amino acid (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Met Gln Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly

Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro

Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg

Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu

Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr

Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu

Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val



# - 109 -

|            |                   |            |            |            |            |            |            |            |             |            | _            |            |            |            |            |
|------------|-------------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|--------------|------------|------------|------------|------------|
| Ar         | g G1              | y Il       | le Pr      | o Ph       | e Tr       | p Gl       | y Gl       | y Se<br>O  | r Th        | r Il       | e As         | p Th<br>12 | r Gl<br>5  | u Le       | u Ly       |
| Va         | 1 I1<br>13        | e As<br>0  | p Th       | r As       | n Cy       | s Il<br>13 | e Ası<br>5 | n Va       | 1 11        | e Gl       | n Pr         | o As<br>O  | p Gl       | y Se       | г Ту:      |
| Ar<br>14   | g Se<br>5         | r Gl       | u Gl       | u Le       | u As:      | n Le       | u Val      | 1 11       | e Il        | e Gl       | y Pro        | o Se       | r Al       | a As       | p Ile      |
| Il         | e Gl              | n Ph       | e Gl       | u Cy<br>16 | s Ly:<br>5 | s Se:      | r Phe      | e Gl       | y Hi.<br>17 | s Gl       | u Vai        | l Le       | u Ası      | n Le       | u Thi      |
| Ar         | g Ası             | n Gl       | y Ty<br>18 | r Gly      | y Sei      | r Thi      | r Glr      | 18!        | r Ile       | e Arg      | g Phe        | s Se       | r Pro      |            | P Ph€      |
| Th:        | r Phe             | 9 Gl       | y Pho      | e Gl       | u Glu      | ı Sei      | 200        | Glu        | ı Va        | l Asp      | Th           | 205        |            | ) Le       | ı Lev      |
| Gly        | y Ala<br>210      | a Gl       | y Ly:      | s Phe      | e Ala      | Thr<br>215 | asp        | Pro        | Ala         | a Val      | . Thr<br>220 |            | a Ala      | His        | Glu        |
| 243        | •                 |            |            |            | . 230      |            |            |            |             | 235        | •            |            |            |            | Asn<br>240 |
|            |                   |            |            | 245        | •          |            |            |            | 250         | )          |              |            |            | 255        |            |
| Glu        | ı Val             | . Sei      | 260        | e Glu      | Glu        | Leu        | Arg        | Thr<br>265 | Phe         | Gly        | Gly          | His        | Asp<br>270 |            | Lys        |
| Phe        | lle               | Asp<br>275 | Ser        | Leu        | Gln        | Glu        | Asn<br>280 | Glu        | Phe         | Arg        | Leu          | Tyr<br>285 | Tyr        | Tyr        | asn        |
| Lys        | Phe<br>290        | Lys        | Asp        | Ile        | Ala        | Ser<br>295 | Thr        | Leu        | Asn         | Lys        | Ala<br>300   | Lys        | Ser        | Ile        | Val        |
| Gly<br>305 | Thr               | Thr        | Ala        | Ser        | Leu<br>310 | Gln        | Tyr        | Met        | Lys         | Asn<br>315 | Val          | Phe        | Lys        | Glu        | Lys<br>320 |
| Tyr        | Leu               | Leu        | Ser        | Glu<br>325 | Asp        | Thr        | Ser        | Gly        | Lys<br>330  | Phe        | Ser          | Val        | Asp        | Lys<br>335 | Leu        |
| Lys        | Phe               | Asp        | Lys<br>340 | Leu        | Tyr        | Lys        | Met        | Leu<br>345 | Thr         | Glu        | Ile          | Tyr        | Thr<br>350 | Glu        | Asp        |
| Asn        | Phe               | Val<br>355 | Lys        | Phe        | Phe        | Lys        | Val<br>360 | Leu        | Asn         | Arg        | Lys          | Thr<br>365 | Tyr        | Leu        | Asn        |
| Phe        | Asp<br>370        | Lys        | Ala        | Val        | Phe        | Lys<br>375 | Ile        | Asn        | Ile         | Val        | Pro<br>380   | Lys        | Val        | Asn        | Tyr        |
| Thr<br>385 | Ile               | Tyr        | Asp        | Gly        | Phe<br>390 | Asn        | Leu        | Arg        | Asn         | Thr<br>395 | Asn          | Leu        | Ala        | Ala        | Asn<br>400 |
| Phe        | Asn               | Gly        | Gln        | Asn<br>405 | Thr        | Glu        | Ile        | Asn        | Asn<br>410  | Met        | Asn          | Phe        | Thr        | Lys<br>415 | Ļeu        |
| Lys        | Asn               | Phe        | Thr<br>420 | Gly        | Leu        | Phe        |            | Phe<br>425 | Tyr         | Lys        | Leu          | Leu        | Cys<br>430 | Val        | Arg        |
| Gly        | Ile               | Ile<br>435 | Thr        | Ser        | Lys        | Thr        | Lys<br>440 | Ser        | Leu         | Asp        |              | Gly<br>445 | Tyr        | Asn        | Lys        |
| Ala        | <i>Leu</i><br>450 | Asn        | Asp        | Leu        | Суз        | Ile<br>455 | Lys        | Val        | Asn         |            | Trp<br>460   | Asp        | Leu        | Phe        | Phe        |
|            |                   |            |            |            |            |            |            |            |             |            |              |            |            |            |            |





- 110 -Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu 485 490 Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu 515 Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu 550 His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys · 580 Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr 615 Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala 630

Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Gleu Phe Ser Gly Ala Val Ile Leu Leu Gleu Phe Ile Pro Glu Ile Ala Gleu Phe Pro Val Leu Gleu Gleu Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys

Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu 690 695 700

680

Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys 705 710 715 720

Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu 725 730 735

Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn 740 745 750

Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp 755 760 765

Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile 770 775 780

Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met 785 790 795 800

Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys 805 810 815



- 111 -

Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly 825

Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp 835 840 845

Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser 850 855

Thr Phe Thr Glu Tyr Ile Lys 870

- (2) INFORMATION FOR SEQ ID NO: 27:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 2574 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: double (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

| ATGCCGGTT  | A CCATCAACA | A CTTCAACTA | C AACGACCCG/ | A TCGACAACA | A CAACATCATC | 60   |
|------------|-------------|-------------|--------------|-------------|--------------|------|
| ATGATGGAA  | C CGCCGTTCG | C ACGTGGTAC | C GGTCGTTACT | C ACAAGGCTT | T CAAGATCACC | 120  |
| GACCGTATC  | I GGATCATCC | C GGAACGTTA | ACCTTCGGTT   | C ACAAACCTG | A GGACTTCAAC | 180  |
| AAGAGTAGC  | GGATTTTCA   | A TCGTGACGT | TGCGAGTACT   | : ATGATCCAG | A TTATCTGAAT | 240  |
|            |             |             |              |             | TAGAATCAAA   | 300  |
|            |             |             |              |             | TTATCTTGGA   | 360  |
|            |             |             | •            |             | TGTTAATAAA   | 420  |
|            |             |             |              |             | TTTAATAATA   |      |
|            |             |             |              |             | ACAAAATCAT   | 480  |
|            |             |             |              |             | AGAATATGTA   | 540  |
|            | ATAATGTTCA  |             |              |             |              | 600  |
|            |             |             |              |             |              | 660  |
|            | CAGCCTTGAT  |             |              |             |              | 720  |
|            | TAGATGATTT  |             |              |             |              | 780  |
|            | TACAGGCAGA  |             |              |             |              | 840  |
|            | CGGATAAAAG  |             |              |             |              | 900  |
| GATAGACTTA | ACAAGGTTTT  | AGTTTGCATA  | TCAGATCCTA   | ACATTAATAT  | TAATATATAT   | 960  |
| TAAAATAAAA | TTAAAGATAA  | ATATAAATTC  | GTTGAAGATT   | CTGAGGGAAA  | ATATAGTATA   | 1020 |
| GATGTAGAAA | GTTTTGATAA  | ATTATATAA   | AGCTTAATGT   | TTGGTTTTAC  | AGAAACTAAT   | 1080 |
| ATAGCAGAAA | ATTATAAAAT  | AAAAACTAGA  | GCTTCTTATT   | TTAGTGATTC  | CTTACCACCA   | 1140 |
| GTAAAAATAA | AAAATTTATT  | AGATAATGAA  | ATCTATACTA   | TAGAGGAAGG  | GTTTAATATA   | 1200 |
|            |             |             |              |             |              |      |



- 112 -

| TCTGATAA  | AG ATATGGAAA  | A AGAATATAG  | A GGTCAGAATA | AAGCTATAAA | TAAACAAGCT | 126    |
|-----------|---------------|--------------|--------------|------------|------------|--------|
| TATGAAGA  | AA TTAGCAAGG  | A GCATTTGGC  | T GTATATAAGA | TACAAATGTG | TAAAAGTGTT | 1320   |
| AAAGCTCC  | AG GAATATGTA  | T TGATGTTGAT | r aatgaagatt | TGTTCTTTAT | AGCTGATAAA | 1386   |
| AATAGTTT  | TT CAGATGATT  | r atctaaaaa  | GAAAGAATAG   | AATATAATAC | ACAGAGTAAT | . 144( |
| TATATAGA  | AA ATGACTTCC  | TATAAATGA    | TTAATTTTAG   | ATACTGATTT | AATAAGTAAA | 1500   |
| ATAGAATTA | AC CAAGTGAAA  | A TACAGAATCA | CTTACTGATT   | TTAATGTAGA | TGTTCCAGTA | 1560   |
|           | AC AACCCGCTAT |              |              |            | . •        | 1620   |
|           | C AGACATTTC   |              |              |            |            | 1680   |
|           | TATTTTCTA     | •            |              |            |            | 1740   |
|           | G TGGTAGAAGO  |              |              |            |            | 1800   |
|           | G AAGCTAATAA  | •            | •            |            |            | 1860   |
| ٠         | A TAGGATTAGO  |              | •            |            |            | 1920   |
|           | G AGATTGCAGG  |              |              |            |            | 1920   |
|           | G GAGCCTTTTT  |              |              |            | •          | 2040   |
|           | A ATGCTTTAAC  |              |              |            | •          |        |
|           | T GGCTCTCAAC  | ·            |              |            |            | 2100   |
|           | A ATTATCAAGC  |              | •            |            | •          | 2160   |
|           | •             |              |              |            |            | 2220   |
|           | A AAGAAAAGTC  |              |              | -          |            | 2280   |
|           | A TTAACCAAGC  |              |              |            |            | 2340   |
|           | A TGAAAAAAT   |              | •            |            |            | 2400   |
|           | A AAAATTTGTT  |              |              |            |            | 2460   |
|           | G AAAAATCAAA  | •            |              |            |            | 2520   |
| CAATATAT  | A CCAATGATAC  | AATACTAATA   | GAAATGTTTA   | ATAAATATAA | TAGC       | 2574   |

#### (2) INFORMATION FOR SEQ ID NO: 28:

- (i) SEQUENCE CHARACTERISTICS:

  (A) LENGTH: 2574 base pairs

  (B) TYPE: nucleic acid

  (C) STRANDEDNESS: double

  (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: DNA (genomic)

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

| 60  | TAATATTATT | TTGATAATAA | AATGATCCTA | TTTTAATTAT | CAATAAATAA | ATGCCAGTTA |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | TAAAATCACA | ATAAAGCTTT | GGGAGATATT | GAGAGGTACG | CTCCATTTGC | ATGATGGAGC |
| 180 | GGATTTTAAT | ATAAACCTGA | ACTTTTGGAT | GGAAAGATAT | GGATAATACC | GATCGTATTT |
| 240 | TTACTTAAAT | ATGATCCAGA | TGTGAATATT | TAGAGATGTT | GTATTTTTAA | AAAAGTTCCG |



## - 113 -

| ACTAATGATA AAAAGAATAT ATTTTTACAA ACAATGATCA AGTTATTTAA TAGAATCAAA | 300  |
|---|------|
| TCAAAACCAT TGGGTGAAAA GTTATTAGAG ATGATTATAA ATGGTATACC TTATCTTGGA | 360  |
| GATAGACGTG TTCCACTCGA AGAGTTTAAC ACAAACATTG CTAGTGTAAC TGTTAATAAA | 420  |
| TTAATCAGTA ATCCAGGAGA AGTGGAGCGA AAAAAAGGTA TTTTCGCAAA TTTAATAATA | 480  |
| TTTGGACCTG GGCCAGTTTT AAATGAAAAT GAGACTATAG ATATAGGTAT ACAAAATCAT | 540  |
| TTTGCATCAA GGGAAGGCTT CGGGGGTATA ATGCAAATGA AGTTTTGCCC AGAATATGTA | 600  |
| AGCGTATTTA ATAATGTTCA AGAAAACAAA GGCGCAAGTA TATTTAATAG ACGTGGATAT | 660  |
| TTTTCAGATC CAGCCTTGAT ATTAATGCAT GAACTCATCC ACGTCCTCCA CGGTCTCTAC | 720  |
| GGTATCAAAG TAGACGACCT CCCGATCGTC CCGAACGAAA AAAAATTCTT CATGCAGAGC | 780  |
| ACCGACGCAA TCCAGGCAGA AGAACTCTAC ACCTTCGGTG GTCAGGACCC GAGCATCATC | 840  |
| ACCCCGAGCA CCGACAAAAG CATCTACGAC AAAGTCCTCC AGAACTTCCG TGGTATCGTC | 90,0 |
| GACCGTCTCA ACAAAGTCCT CGTCTGCATC AGCGACCCGA ACATCAACAT CAACATCTAC | 960  |
| AAAAACAAAT TCAAAGACAA ATACAAATTC GTCGAAGACA GCGAAGGTAA ATACAGCATC | 1020 |
| GACGTCGAGA GCTTCGACAA ACTCTACAAA AGCCTCATGT TCGGTTTCAC CGAAACCAAC | 1080 |
| ATCGCAGAAA ACTACAAAAT CAAAACCCGT GCAAGCTACT TCAGCGACAG CCTCCCGCCG | 1140 |
| GTCAAAATCA AAAACCTCCT CGACAACGAA ATCTACACCA TCGAAGAAGG TTTCAACATC | 1200 |
| AGCGACAAAG ACATGGAAAA AGAATACCGT GGTCAGAACA AAGCAATCAA CAAACAAGCT | 1260 |
| TACGAAGAA TCAGCAAAGA ACACCTCGCA GTCTACAAAA TCCAGATGTG CAAAAGCGTC  | 1320 |
| AAAGCACCGG GTATCTGCAT CGACGTTGAC AACGAAGACC TCTTCTTCAT CGCAGACAAA | 1380 |
| AACAGCTTCA GCGACGACCT CAGCAAAAAC GAACGTATCG AATACAACAC CCAGAGCAAC | 1440 |
| TACATCGAAA ACGACTTCCC GATCAACGAA CTCATCCTCG ACACCGACCT CATCAGCAAA | 1500 |
| ATCGAACTCC CGAGCGAAAA CACCGAAAGC CTCACCGACT TCAACGTTGA CGTCCCGGTC | 1560 |
| TACGAAAAAC AGCCGGCAAT CAAAAAAATC TTCACCGACG AAAACACCAT CTTCCAGTAC | 1620 |
| CTCTACAGCC AGACCTTCCC GCTAGATATA AGAGATATAA GTTTAACATC TTCATTTGAT | 1680 |
| GATGCATTAT TATTTTCTAA CAAAGTTTAT TCATTTTTTT CTATGGATTA TATTAAAACT | 1740 |
| GCTAATAAAG TGGTAGAAGC AGGATTATIT GCAGGTTGGG TGAAACAGAT AGTAAATGAT | 1800 |
| TTTGTAATCG AAGCTAATAA AAGCAATACT ATGGATAAAA TTGCAGATAT ATCTCTAATT | 1860 |
| GTTCCTTATA TAGGATTAGC TTTAAATGTA GGAAATGAAA CAGCTAAAGG AAATTTTGAA | 1920 |
| AATGCTTTTG AGATTGCAGG AGCCAGTATT CTACTAGAAT TTATACCAGA ACTTTTAATA | 1980 |
| CCTGTAGTTG GAGCCTTTTT ATTAGAATCA TATATTGACA ATAAAAATAA AATTATTAAA | 2040 |
| ACAATAGATA ATGCTTTAAC TAAAAGAAAT GAAAAATGGA GTGATATGTA CGGATTAATA | 2100 |
| GTAGCGCAAT GGCTCTCAAC AGTTAATACT CAATTTTATA CAATAAAAGA GGGAATGTAT | 2160 |
| AAGGCTTTAA ATTATCAAGC ACAAGCATTG GAAGAAATAA TAAAATACAG ATATAATATA | 2220 |
| TATTCTGAAA AAGAAAAGTC AAATATTAAC ATCGATTTTA ATGATATAAA TTCTAAACTT | 2280 |
|   |      |

#### - 114 -

| AATGAGGGTA | TTAACCAAGC | TATAGATAAT | ATAAATAATT | TTATAAATGG | ATGTTCTGTA |      |
|------------|------------|------------|------------|------------|------------|------|
| TCATATTTAA | TGAAAAAAAT | GATTCCATTA | GCTGTAGAAA | AATTACTAGA | CTTTGATAAT | 2340 |
| ACTCTCAAAA | AAAATTTGTT | AAATTATATA | GATGAAAATA | AATTATATTT | GATTGGAAGT | 2400 |
| GCAGAATATG | AAAAATCAAA | AGTAAATAAA | TACTTGAAAA | CCATTATGCC | GTTTGATCTT | 2460 |
| TCAATATATA | CCAATGATAC | AATACTAATA | GAAATGTTTA | ATAAATATAA | TAGC       | 2520 |
| •          | •          |            |            |            |            | 2574 |





## **CLAIMS**

- 1. A polypeptide comprising first and second domains, wherein said first domain is adapted to cleave one or more vesicle or plasma-membrane associated proteins essential to exocytosis, and wherein said second domain is adapted (i) to translocate the polypeptide into a cell or (ii) to increase the solubility of the polypeptide compared to the solubility of the first domain on its own or (iii) both to translocate the polypeptide into a cell and to increase the solubility of the polypeptide compared to the solubility of the first domain on its own, said polypeptide being free of clostridial neurotoxin and free of clostridial neurotoxin precursor that can be converted into toxin by proteolytic action.
- 2. A polypeptide according to Claim 1 wherein said first domain comprises a clostridial toxin light chain.
- 3. A polypeptide according to Claim 1 wherein said first domain comprises a fragment or variant of a clostridial toxin light chain.
- 4. A polypeptide according to Claim 2 or 3 wherein the clostridial toxin is a botulinum toxin.
- 5. A polypeptide according to any preceding claim wherein the first domain exhibits endopeptidase activity specific for a substrate selected from one or more of SNAP-25, synaptobrevin/VAMP and syntaxin.
- 6. A polypeptide according to any preceding claim wherein said second domain comprises a clostridial toxin heavy chain  $H_N$  portion.
- 7. A polypeptide according to any of Claims 1-5 wherein said second domain comprises a fragment or variant of a clostridial toxin heavy chain  $H_N$  portion.
- 8. A polypeptide according to Claim 6 or 7 wherein the clostridial toxin is a

- 116 -

#### botulinum toxin.

- 9. A polypeptide according to any of Claims 1-8 further comprising a third domain adapted for binding of the polypeptide to a cell, by binding of the third domain directly to a cell or by binding of the third domain to a ligand or to ligands that bind to a cell.
- 10. A polypeptide according to Claim 9 wherein said third domain is for binding the polypeptide to an immunoglobulin.
- 11. A polypeptide according to Claim 10 wherein said third domain is a tandem repeat synthetic IgG binding domain derived from domain  $\beta$  of Staphylococcal protein A.
- 12. A polypeptide according to Claim 9 wherein said third domain comprises an amino acid sequence that binds to a cell surface receptor.
- 13. A polypeptide according to Claim 12 wherein said third domain is insulin-like growth factor-1 (IGF-1).
- 14. A polypeptide according to any preceding claim comprising a botulinum toxin light chain or a fragment or a variant of a botulinum toxin light chain and a portion designated  $H_N$  of a botulinum toxin heavy chain.
- 15. A polypeptide according to Claim 14 wherein one or both of (a) the toxin light chain or fragment or variant of toxin light chain and (b) the portion of the toxin heavy chain are of botulinum toxin type A.
- 16. A polypeptide according to Claim 15 wherein the botulinum toxin type A light chain variant has at residue 2 a glutamate, at residue 26 a lysine and at r sidue 27 a tyrosine.

- 117 -

- 17. A polypeptide according to Claim 14 wherein one or both of (a) the toxin light chain or fragment or variant of toxin light chain and (b) the portion of the toxin heavy chain are of botulinum toxin type B.
- 18. A polypeptide according to any of Claims 1-13 comprising a botulinum toxin light chain or a fragment or a variant of a botulinum toxin light chain and at least 100 N-terminal amino acids of a botulinum toxin heavy chain.
- 19. A polypeptide according to Claim 18 comprising a botulinum toxin type B light chain, or a fragment or variant thereof, and 107 N-terminal amino acids of a botulinum toxin type B heavy chain.
- 20. A polypeptide according to Claim 15 or 16 comprising at least 423 of the N-terminal amino acids of botulinum toxin type A heavy chain.
- 21. A polypeptide according to Claim 20 comprising a botulinum toxin type A light chain and 423 N-terminal amino acids of a botulinum toxin type A heavy chain.
- 22. A polypeptide according to Claim 20 comprising a botulinum toxin type A light chain variant wherein residue 2 is a glutamate, residue 26 is a lysine and residue 27 is a tyrosine, and 423 N-terminal amino acids of a botulinum toxin type A heavy chain.
- 23. A polypeptide according to Claim 17 comprising at least 417 of the N-terminal amino acids of botulinum toxin type B heavy chain.
- 24. A polypeptide according to Claim 23 comprising a botulinum toxin type B light chain and 417 N-terminal amino acids of a botulinum toxin type B heavy chain.
- 25. A polypeptide according to any of Claims 14-24 lacking a portion designated

H<sub>c</sub> of a botulinum toxin heavy chain.

- 26. A polypeptide comprising a botulinum toxin light chain and a fragment of a botulinum toxin heavy chain, said fragment being not capable of binding to cell surface receptors.
- 27. A polypeptide according to Claim 26 lacking an intact portion designated  $H_{\text{c}}$  of a botulinum toxin heavy chain.
- 28. A polypeptide according to any preceding claim comprising a variant of a clostridial toxin and further comprising a site for cleavage by a proteolytic enzyme, which cleavage site is not present in the native toxin.
- 29. A polypeptide according to Claim 28 comprising a variant of a clostridial toxin light chain and further comprising a site for cleavage by a proteolytic enzyme, which cleavage site is not present in the native toxin light chain.
- 30. A polypeptide according to Claim 28 or 29 comprising a variant of a clostridial toxin heavy chain  $H_N$  portion and further comprising a site for cleavage by a proteolytic enzyme, which cleavage site is not present in the native toxin heavy chain  $H_N$  portion.
- 31. A polypeptide according to Claim 28, 29 or 30 obtainable by modification of a DNA encoding the polypeptide so as to introduce one or more nucleotides coding for the cleavage site.
- 32. A fusion protein comprising a fusion of (a) a polypeptide according to any of Claims 1-31 with (b) a second polypeptide being a polypeptide or oligopeptide adapted for binding to an affinity matrix so as to enable purification of the fusion protein using said matrix.
- 33. A fusion protein according to Claim 32 wh rein said second polypeptide is



adapted to bind to a chromatography column, such as an affinity matrix of glutathione Sepharose.

- 119 -

- 34. A fusion protein according to Claim 32 or 33 wherein a specific protease cleavage site is incorporated between the first and second polypeptides, said protease site enabling proteolytic separation of first and second polypeptides.
- 35. A composition comprising a derivative of a clostridial toxin, said derivative retaining at least 10% of the endopeptidase activity of the botulinum toxin, said derivative further being non-toxic *in vivo* due to its inability to bind to cell surface receptors, and wherein the composition is free of any component, such as toxin or a further toxin derivative, that is toxic *in vivo*.
- 36. A composition according to Claim 35 or a polypeptide according to any of Claims 1-31 or a fusion protein according to Claim 32, 33 or 34 for use as a positive control in a toxin assay.
- 37. A composition according to Claim 35 or a polypeptide according to any of Claims 1-31 or a fusion protein according to Claim 32, 33 or 34 for use as a vaccine against clostridial toxin.
- 38. A composition according to Claim 35 or a polypeptide according to any of Claims 1-31 or a fusion protein according to Claim 32, 33 or 34 for *in vivo* use.
- 39. A pharmaceutical composition comprising a composition according to Claim 35, a polypeptide according to any of claims 1-31 or a fusion protein according to Claim 32, 33 or 34, in combination with a pharmaceutically acceptable carrier.
- 40. A nucleic acid encoding a polypeptide or a fusion protein according to any of Claims 1-34.
- 41. A nucleic acid encoding a polypeptide or a fusion protein according to Claim



- 120 -

40 and comprising nucleotides encoding residu s 1-448 of a botulinum toxin type A light chain.

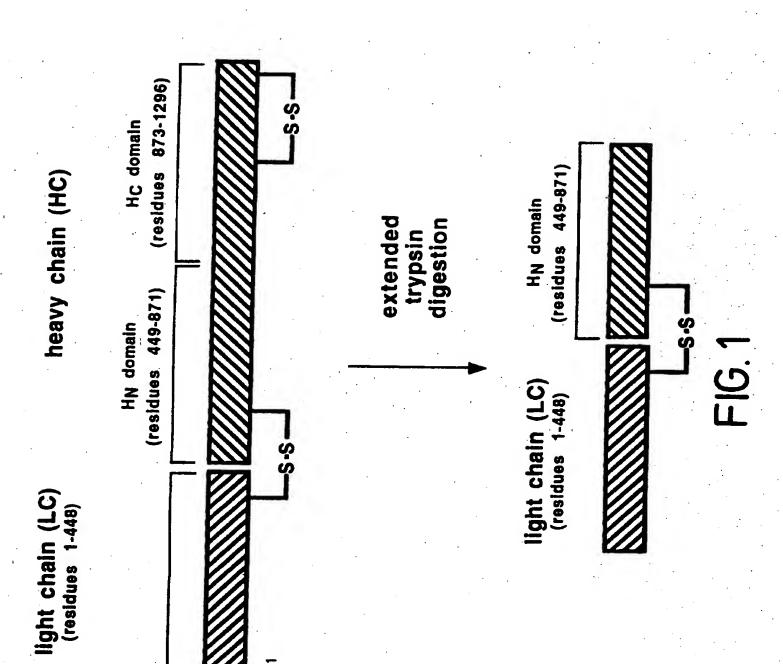
- 42. A nucleic acid according to Claim 40 or 41 comprising nucleotides encoding residues 1-423 of a botulinum toxin type A heavy chain  $H_N$  domain.
- 43. A nucleic acid encoding a polypeptide or a fusion protein according to Claim 40 and comprising nucleotides encoding residues 1-470 of a botulinum toxin type B light chain.
- 44. A nucleic acid encoding a polypeptide or a fusion protein according to Claim 40 or 43 comprising nucleotides encoding residues 1-417 of a botulinum toxin type B heavy chain  $H_N$  domain.
- 45. A nucleic acid according to any of Claims 40-44 comprising nucleotides encoding a restriction endonuclease cleavage site not present in native clostridial toxin sequence.
- 46. A nucleotide according to Claim 45 obtainable by modification of a nucleotide encoding a polypeptide or fusion protein according to any of claims 1-34 so as to introduce said cleavage site.
- 47. A DNA according to any of claims 40-46.
- 48. A DNA selected from SEQ ID No:s 1, 8, 10, 12, 14, 16, 18, 23 and 24.
- 49. A method of manufacture of a polypeptide according to any of Claims 1-31 comprising expressing in a host cell a nucleic acid according to any of Claims 40-48 and recovering the polypeptide.
- 50. A method of manufacture of a polypeptide according to any of Claims 1-31 comprising expressing in a host cell a nucleic acid encoding a fusion protein



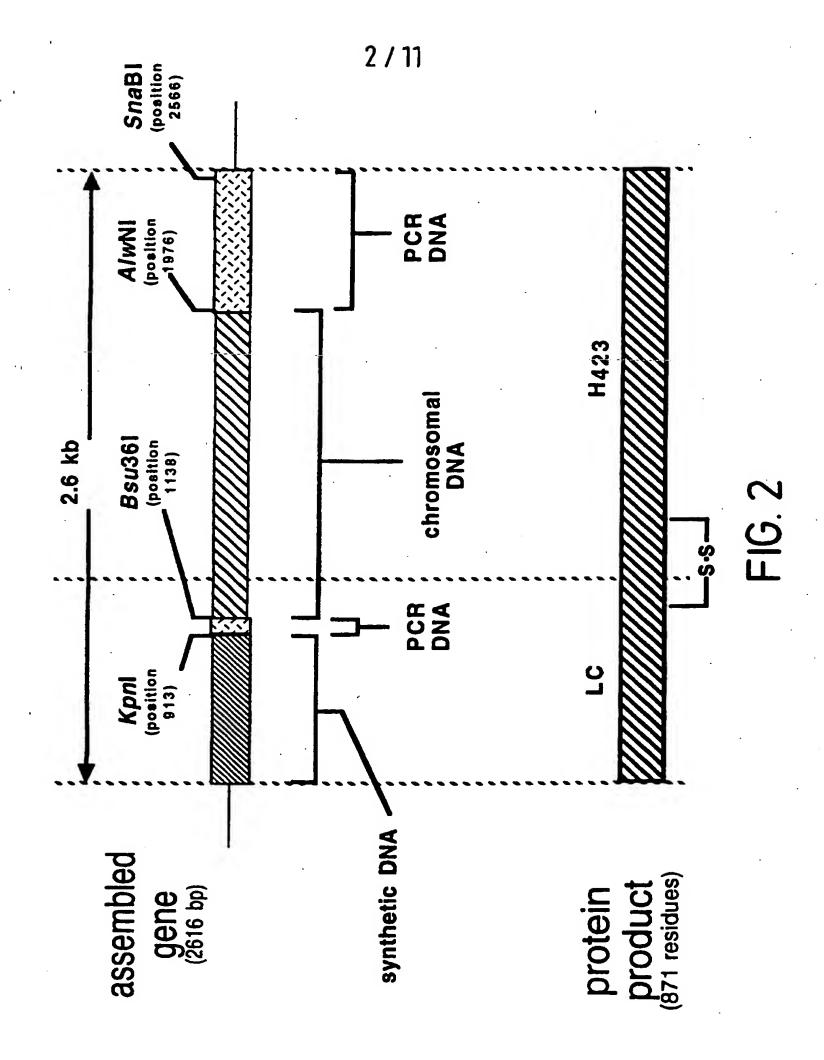
according to Claim 32, 33 or 34, purifying the fusion protein by eluting the fusion protein through an affinity matrix adapted to retain the fusion protein and eluting through said matrix a ligand adapted to displace the fusion protein, and recovering the fusion protein.

- A method of manufacture according to Claims 49 or 50 in which the nucleic 51. acid is DNA.
- A cell expressing a polypeptide or fusion protein according to any of Claims **52**. 1-34.

1/11







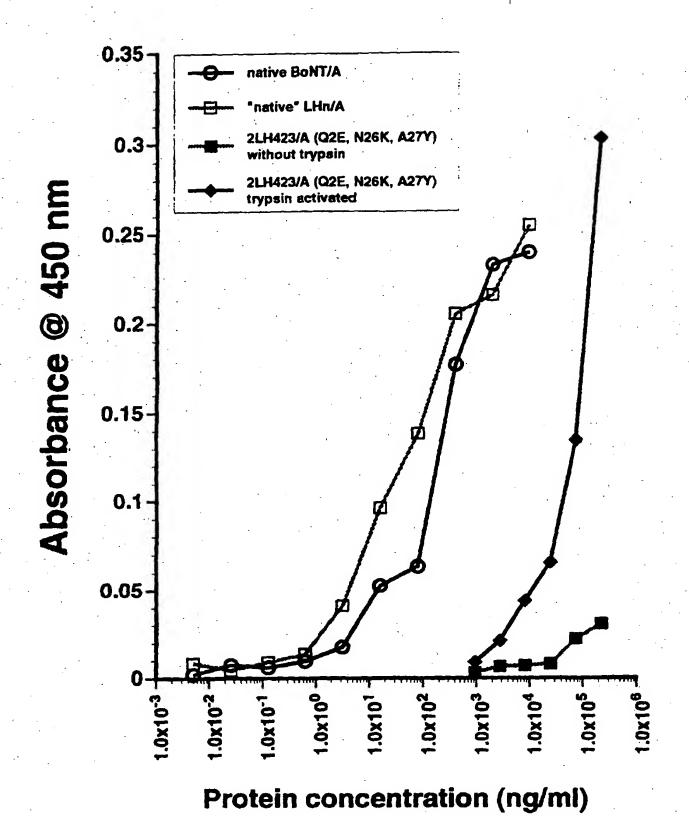


FIG. 3

|          | 4 |
|----------|---|
| •        | 4 |
| <b>T</b> |   |
|          | 7 |

|   |                      |                         | 4/1                                  | 1  |   |
|---|----------------------|-------------------------|--------------------------------------|--|---|
| Q P V (Seq 1.D. 2)  | •                    | (Seq I.D. 4)            | ). 6)                                |  |   |
| <u>5</u>  |                      | ).<br> -<br> -          | (Seq I.D.                            |  |   |
| (Se   |                      |                         | S)                                   |  |   |
| >   |                      | > .                     | >                                    | >  | >   |
| <u>a</u> .  |                      | هـ                      | <u>م</u>                             | <b>Q</b> .   | <u> </u>  |
| -   |                      | 0                       | 0                                    | . 0  | 0   |
| ~   |                      | <b>∑</b> .              | <b>∑</b>                             | ∑ ⊘  | Σ   |
| Σ<br>Ο<br><u>υ</u>  |                      | <u> </u>                | ک<br>و                               | ့်မ  | ۵<br>و  |
| Image: Control of the |                      | <u> </u>                | <u>→</u>                             | <u> </u>   | <b>▼</b>  |
| K I P N A   |                      | ×                       | ×                                    | Z  | z   |
| 4   |                      | ۵                       | ۵                                    | Δ.   | ۵   |
| H   |                      | H                       | H                                    | H  | <b></b>   |
|   |                      | ¥                       | ×                                    | ×  | ¥   |
| <b>H</b> .  |                      | H                       | H                                    | H  | H   |
| <b>≻</b>  |                      | <b>≻</b> .              | <b>≻</b>                             | <b>&gt;</b>  | <b>*</b>  |
| 7 1   | Δ.                   | Į į                     | Į                                    | H  | ı,  |
| D I   | _                    | _                       | _<br>_                               | <u> </u>   |   |
| >   | ш                    | >                       | >                                    | >  | >   |
| <b>5</b>  | ıL                   | G                       | <b>ن</b> .                           | ی  | 5   |
| Z   | ш                    | 2                       | Z                                    | Z  | Z   |
| >   |                      | >                       | >                                    | >  | >   |
| •   | _                    | <b>Q</b> .              | ۵.                                   | •  | ٥   |
| 0   | ×                    | ۵                       | 9                                    | 9  | 0   |
| <b>⊻</b> .  | 0                    | <b>~</b>                | <b>.</b>                             | <b>×</b>   | <b>×</b>  |
| <b>&gt;</b>   | <del></del>          | <u>~</u>                | <b>&gt;</b>                          | _  | <b>-</b>  |
| 2   |                      | <u> </u>                | 2                                    | <b>Z</b>   | 2   |
| 8   |                      | ~                       | ~                                    | ~  | ~   |
| <b>Y</b>  | <b>-</b>             | ×                       | ~                                    | ×  | ¥   |
| Z   | X                    | Z                       | Z                                    | Z  | Z   |
| >   | I                    | >                       | >                                    | >  | >   |
|   | H                    | <u>u</u>                | <u>u</u>                             | u.   | <u> </u>  |
| MQFVNKQFNYKDP<br>1  | GSPGIHMTSTRLQKL<br>1 | GTMEFVNKQFNYKDP<br>- 24 | G S M E F V N K Q F N Y K D P<br>123 | M Q F V N K Q F N Y K D P                                  | MPFVNKQFNYKDP<br>1                                    |
| <b>₹</b> ←  | •                    | - M                     | x w                                  | <b>X</b> =   | <b>≖</b> ⊣  |
|   | S                    | <u> </u>                | 2 .                                  |  |   |
|   | <u> </u>             |                         | 0 4                                  |  |   |
|   |                      | 427Y)                   | A27Y)                                | 169<br>a/.1990   | A C   |
| LH <sub>423</sub> /A  | 23LH423/A            | (Q2E, N26K, A27Y)       | 2LH423/A<br>(Q2E, N26K, A27Y)        | Native BoNT/A,<br>C. botulinum 2169<br>Thompson et al.1990 | Native BoNT/A,<br>C. botulinum 62A<br>Binz et al.1990 |
|   | .,                   |                         |                                      | Ž O Ĕ  | ž C 🛱   |

= REGIONS OF NON-IDENTITY WITH THE NATIVE SEQUENCES.

FIG 4

1321/441

TCA TTA GAT AAA GGA TAC AAT AAG agc gct gat ggg GCA TTA AAT GAT TTA TGT ATC AAA

TCA TTA GAT AAA GGA TAC AAT AAG agc gct gat ggg GCA TTA AAT GAT TTA TGT ATC AAA

S L D K G Y N K S A D G A L N D L C I K
S L D K G Y N K S A D G A L N D L C I K H-chain

F16.5

1321/441 TCA TTA GAT AAA GGA TAC AAT AAG atc gaa ggt cgt tgc gat ggg GCA TTA AAT GAT TTA S L D K G Y N K I E G R C D G A L N D L H-chain Factor Xa protease motif



6/11

CAG ACA Q T GAG ATG . GGA G GCG A **T**6T AAG K GTG V AGG R CAG TTC Q F ر 30 AGG AGG R AGC S CGG AGC GGG TAT G Y CT6 CCC ACA GAG E S S S AAG K 999 9 167 C GAG E AAG GAT 2587/863 TAC GTA GAT / Y V D I 2647/883 CCG GAG ACG P P E T 2707/903 GGC TTT TAT G F Y G F Y G F Y G F Y G F Y G F Y

=16.7

ACA CAT AAT AGG R GAT GAT D TAT Y 2617/873

\_A TTT ACT GAA T'

f F T E
2677/893

GGC TAT CAA T' AGA R **₹**~ S & AAA K E SA A N GAT D 2587/863 TAC GTA (Y Y V [ 2647/883 TCT AAA ( S K V

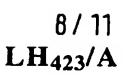
F1G. 8

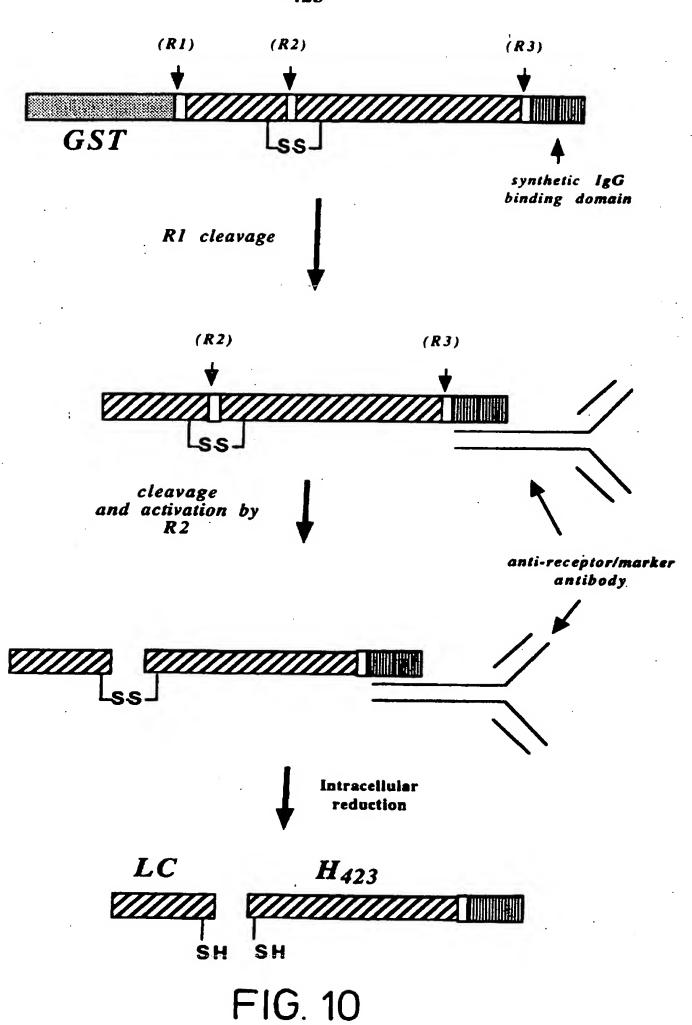


7/11

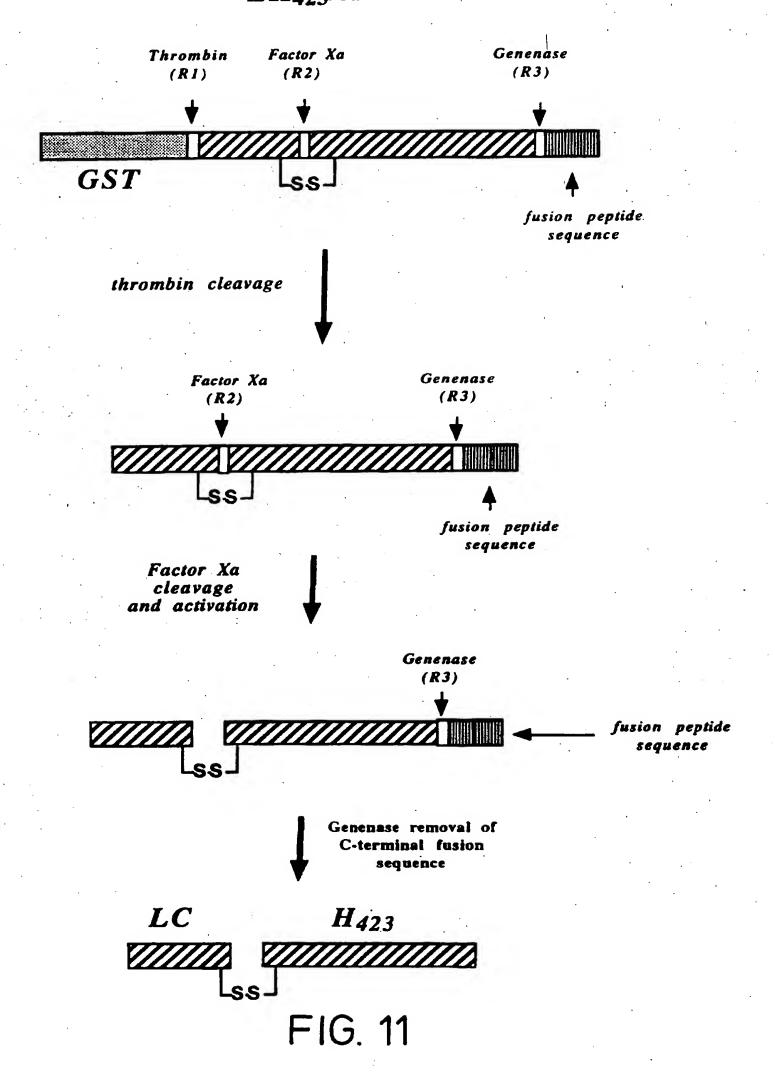


AAA K S S GAA E AAA K R SA AAT N A A GAA E AAC N AAC N A C CTG L CGA R AAA K TTA ၂၂ ၂၂ ၂၂ GAA E CAA O TCA AAC GAA E AAA K GAA E AAG AAC N GAA E SCT AAC N AGC GAC D CAA 2977/993 CAA AGC 0 S 7/933 AGC S 5/7953 CTA GAA E 7913 2917 <u>ي</u> درو OAC D ACA E CAC GCG A ور م ا GAC D GAG E 0 CAG TA L ATC I GAT GCT A TAT Y TAT S S GAG E AAA K E i GAT D AGT S E ST TAT Y 80 \$ 0 AAT N SCT A S S A AAC I Q S 3007/1003 AG CT 827/943 3CT AAA AAG C A K K 2887/963 CAA AAC GCC Q N A 2947/983 TTC ATC C SGT CAA 



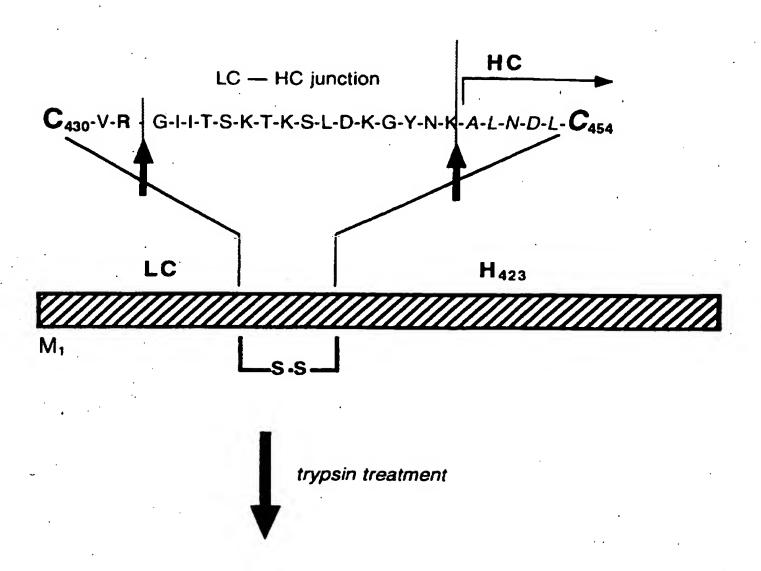


# $LH_{423}/A^{9/11}$









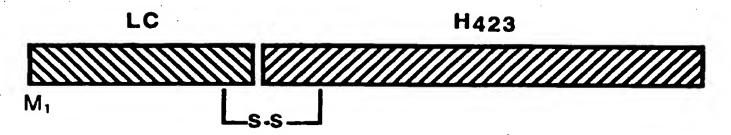
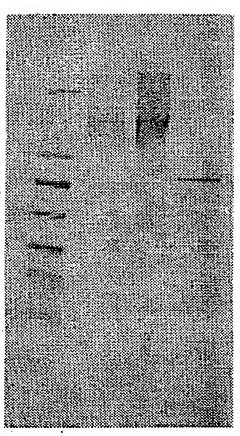


FIG. 12



11 / 11

Panel A.
1 2 3 4



Panel B. 1 2 3 4

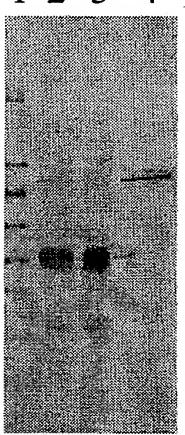
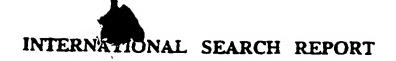


FIG. 13





A. CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/31 C12N1/21 C12P21/02 C07K14/33 A61K38/16 A61K39/08 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N C12P A61K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Category \* Relevant to claim No. X WO 96 12802 A (OPHIDIAN PHARM INC 1-52 ; WILLIAMS JAMES A (US); PADHYE NISHA V (US); KI) 2 May 1996 see the whole document KURAZONO H ET AL: "Minimal essential X 1-52 \*domains\* specifying toxicity of the \*light\* \*chains\* of tetanus toxin and botulinum neurotoxin type A." J BIOL CHEM, JUL 25 1992, 267 (21) P14721-9, UNITED STATES, XP002047910 see table II -/--Further documents are fisted in the continuation of box C. Patent family members are listed in annex. X \* Special outagories of cited documents : \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such docu-ments, such combination being obvious to a person skilled in the art. "O" document referring to an oral disclosure, use, exhibition or socument published prior to the international filing date but tater than the priority date olaimed \*&\* document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 3 0. OL 98 9 December 1997 Name and mailing address of the ISA **Authorized officer** European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Ripsrijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016 Hillenbrand, G



|                | <u> </u>       |
|----------------|----------------|
| Interior which | Application No |
| PCT/GB         | 97/02273       |

| •           |  | PC1/GB 9//022/3       |
|-------------|--|-----------------------|
| C.(Continue | tion) DOCUMENTS CONSIDERED TO BE RELEVANT  |                       |
| Category *  | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
| X           | LI ET AL: "Á SINGLE MUTATION IN THE<br>RECOMBINANT LIGHT CHAIN OF TETANUS TOXIN<br>ABOLISHES ITS PROTEOLYTIC ACTIVITY AND  | 1                     |
|             | REMOVES THE TOXICITY SEEN AFTER RECONSTITUTION WITH NATIVE HEAVY CHAIN" BIOCHEMISTRY, vol. 33, no. 22, 1994,   | *                     |
|             | pages 7014-7020, XP002015938 see the whole document  | 1 26 25               |
| <b>A</b>    | BINZ T ET AL: "THE COMPLETE SEQUENCE OF BOTULINUM NEUROTOXIN TYPE A AND COMPARISON WITH OTHER CLOSTRIDIAL NEUROTOXINS" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 265, no. 16, 5 June 1990, pages 9153-9158, XP002009348 see the whole document | 1,26,35               |
|             |  |                       |
|             |  |                       |
|             |  |                       |
|             |  |                       |
|             |  |                       |
|             |  |                       |
|             |  |                       |
|             |  |                       |
|             |  |                       |



# INTERNATIONAL SEARCH REPORT

Intel. .onal Application No PCT/GB 97/02273

| Patent document cited in search report | Publication<br>date | Patent family member(s) |     | Publication date          |
|--|---------------------|-------------------------|-----|---------------------------|
| WO 9612802 A                           | 02-05-96            | US 5601823 /            | Α , | 11-02-97                  |
|  |                     | US 5196193 A            | A 🦠 | 23-03-93                  |
|  |                     | AU 3968395 /            | A ' | 15-05-96                  |
|  |                     | EP 0796326 /            | A   | 24-09-97                  |
|  |                     | FI 971732 /             | A   | 23 <b>-</b> 06 <b>-97</b> |
|  |                     | NO 971868 /             | A   | 24-06-97                  |
|  |                     | PL 320214 /             | Ą   | 15-09 <b>-</b> 97         |
|  |                     | ZA 9508990 A            | A   | 15-05-96                  |
|  |                     | AU 6653894 A            | A   | 04 <b>-07-94</b>          |
|  |                     | CA 2150935 A            | 4   | 23-06-94                  |
|  | • •                 | EP 0671902 A            | 4   | 20-09-95                  |
|  |                     | WO 9413264 A            | 4   | 23-06-94                  |
|  |                     | US 5466672 A            | 1   | 14-11-95                  |
|  |                     | US 5599539 A            | 1   | 04-02-97                  |
|  | •                   | AU 638786 B             | 3   | <b>08-07-93</b>           |
|  |                     | AU <u>689519</u> 1 A    | 1   | 31-05- <u>91</u>          |
|  | •                   | EP 0498854 A            | 1   | 19-08-92                  |
| ,                                      |                     | WO 9106306 A            | ١   | 16-05-91                  |
|  |                     | US 5443976 A            | 1   | 22-08-95                  |
|  |                     | US 5340923 A            |     | 23-08-94                  |

THIS PAGE BLANK (USPTO)